Physics 5C: Introductory Thermodynamics and Quantum Mechanics **Fall 2018**

Lecture 1: Introduction, Thermodynamics and Combinatorics

Lecturer: Feng Wang *23 August* *Aditya Sengupta*

Note: LaTeX*format adapted from template courtesy of UC Berkeley EECS dept.*

1.1 Overview

There are two major ways of approaching thermodynamics. The first is a phenomenological view, based on looking at macroscopic systems. These systems are specified by various quantities, such as temperature, energy, work, and pressure. Observing these led to the first and second laws of thermodynamics, as well as the idea of entropy. The second is a microscopic view; the study of particle behaviour to explain the macroscopic effects. This is the basic idea behind *statistical mechanics*, which treats the actions of particles as essentially random and therefore, on large scales, predictable (random variations average out over a large number of particles).

This course will start with a discussion of concepts in thermal physics such as temperature, as well as an introduction to probability, then it will cover the macroscopic and microscopic views of thermodynamics described above.

1.2 Large Number Systems

Microscopically, the objects we are considering are atoms and molecules, on size scales of about $10^{-10}m$, and macroscopically, the objects are more on the scale of meters and centimeters. This is a considerable gap. Consider a cube with side length $1cm$. The number of atoms of size $10^{-10}cm$ that can fit in it is

$$\left(\frac{10^{-2}m}{10^{-10}m}\right)^3 = 10^{24} \tag{1.1}$$

The actual quantity that acts as a bridge between macro and micro systems is not far off from this. That is Avogadro's number, $N_A \approx 6.022 \times 10^{23}$. It is defined as the number of carbon atoms in 12g of ^{12}C.

An easy unit to use in macroscopic systems is the mole. One mole of something is defined as the quantity of matter that contains N_A objects (atoms or molecules). For example, one mole of hydrogen is the quantity required to have 6.022×10^{23} atoms of hydrogen.

The macroscopic view of thermodynamics is provably correct because it is within the so-called *thermodynamic limit* of the number of particles, the point after which the random nature of each particle mostly averages out. It is formally defined as a ratio of a number of particles to the volume containing them, but it can be easily understood conceptually.

Consider a system consisting of one particle and a pressure sensor on one wall that indicates when a particle collides with it. This system would exhibit sporadic behavior as the sensor is intermittently triggered randomly, because the particle that can collide with it moves randomly.

Now, consider a similar system with 20 particles. Now, the sensor may average a pressure of 2 or 3 collisions per Δt time interval, but it will still fluctuate, albeit not quite as much.

Increase the number of particles to 10^6, and the pressure begins to converge. There are still fluctuations, but they are on a much smaller scale; now, the number of collisions averages maybe 100,000. Although it may dip to 99,700 or rise to 100,105, the relative fluctuations will be much smaller.

As the number of particles N increases, the absolute fluctuation increases linearly with N, but the relative fluctuation scales only with $\sqrt{N}$, to the point that when we reach a macroscopic-sized system of N_A particles, the force felt by the sensor is practically a constant.

1.3 Macroscopic Parameters

This allows us to define pressure as the ratio of force to the area over which it is applied. Pressure is a *thermodynamic state variable*, meaning that it is an averaged physical quantity that can be used to describe a macroscopic system in a useful way. The state of a system can be completely defined in terms of state variables such as pressure and temperature.

We'll now try and select other state variables to give a useful description of a macroscopic system. The averaged velocity and momentum, $\langle\vec{v}\rangle$ and $\langle\vec{p}\rangle$, are both 0. This is because it is a static system, in which no one direction is preferred over any other. The average kinetic energy can, however, be well-defined and is potentially useful. The quantity that is actually used is U, the total energy of a particle. This is the most important quantity in macro-analysis.

Already, with only a vague overview of how state variables define a system, we can use them to understand what will happen to a system under some circumstances. Let there be a box with known pressure, volume, cumulative particle energy, and temperature, which we then split in half. We then want to predict these state variables for one of the smaller boxes thus created. Clearly, the volume is halved, as is the cumulative particle energy because the number of particles is halved. These are called *extensive state variables* because they depend (usually linearly) on system size. But the temperature remains the same (this is more intuitive but can be more thoroughly shown later) and the pressure is also the same because the number of particles exerting a force is halved, and the area over which that force is distributed is also halved. These two are therefore *intensive state variables.*

1.4 Ideal Gases

We want to find relations between U, P, V, and T that let us predict system behavior in dynamic systems, so that we can analyze cases beyond the above. Consider a system that has a fixed P, V, and T. To experimentally find relations, we can fix one and vary the other two to find a correlation.

- If we fix T, we find an inverse relationship between P and V. As volume increases, pressure reduces because of a lower probability of collisions. We find that

  $$P \propto \frac{1}{V} \tag{1.2}$$

  This is called *Boyle's law.*

- If we fix P, we find that temperature and volume increase together. This can be understood as a formal representation of heat expansion, which we intuitively understand. Therefore

  $$V \propto T \tag{1.3}$$

  This is *Charles' law.*

- If we fix V, we find that pressure and tempeature increase together.

  $$P \propto T \tag{1.4}$$

  This is *Gay-Lussac's law.*

We can combine these into one equation, along with the intuitive-but-will-later-be-rigorously-shown fact that N scales proportional to V, to make the *Ideal Gas Law*:

$$PV \propto NT \tag{1.5}$$

$$PV = Nk_BT \tag{1.6}$$

k_B here is the *Boltzmann constant.*

A more natural description of a macroscopic system would involve a number of moles, so that we don't have to deal with two mystery constants; k_B and N_A. Therefore the Ideal Gas Law can be better stated as follows:

$$PV = \frac{N}{N_A}(N_A k_B)T \tag{1.7}$$

$$PV = nRT \tag{1.8}$$

In using this equation, we are assuming the following reasonably good approximations:

- the size of a particle is negligible
- there is no molecule-molecule interaction

1.5 Combinatorial Problems

A lot of the microscopic view of thermodynamics will involve simply counting particles and combinations thereof. Combinatorics can yield numbers much greater than N_A, so it's important to build techniques to deal with combinations.

1.5.1 Combinations

In a system of n atoms, each one has energy 0 or 1. The system has a total energy of r. How many distinct arrangements are possible?

In the system, there will be r atoms with energy 1, and $n-r$ atoms will have energy 0. To begin with, let there be n distinct numbers $(i_1, i_2, \ldots, i_n)$ and n spaces where they can fit. For the first space, there are n choices since any of the i_ks can go in the first space. Once the first space has been filled, there are $n-1$ choices left for the second space, $n-2$ for the third, and so on, until there is only one choice for the last space. Therefore, the total number of choices for overall arrangements of the system if all the numbers are distinct is

$$n \times (n-1) \times (n-2) \times \cdots \times 2 \times 1 := n! \tag{1.9}$$

What if they are not all distinct? In our case, r of the numbers are 1 and $n-r$ are 0. To resolve this, consider a case in which there are m spaces but all the numbers are the same. We saw above that the all-distinct treatment yields $m!$ as the number of distinct arrangements, but there is only 1 distinct arrangement because all $m!$ arrangements are the same. So we can divide by $m!$ to correct for double-counting.

In our case, we are double-counting the r 1s, so we divide by $r!$. We are also double-counting the $n-r$ 0s, so we divide by $(n-r)!$. This gives us the following formula:

$$choices = \frac{n!}{r!(n-r)!} = {}^nC_r = \binom{n}{r} \tag{1.10}$$

1.5.2 Stirling's Formula

Factorials grow quickly. For even a simple system consisting of 100 atoms, 40 of which have an energy of 1, $\binom{n}{r} \approx 10^{28}$. When n grows to become close to N_A, this is unreasonably large.

To simplify, we use Stirling's formula:

$$\ln n! = n \ln n - n + \ldots \tag{1.11}$$

This helps us make sense of massive factorials. For example, $10^{23}!$:

$$\ln(10^{23}!) = 10^{23} \ln 10^{23} - 10^{23} \tag{1.12}$$

$$\ln(10^{23}!) = 10^{23} \left(23 \ln 10 - 1\right) \tag{1.13}$$

$$\ln(10^{23}!) = 52 \times 10^{23} \tag{1.14}$$

$$\therefore 10^{23}! = e^{52 \times 10^{23}} = 10^{\log_{10} e \times 52 \times 10^{23}} \tag{1.15}$$

$$10^{23}! \approx 10^{2.26 \times 10^{24}} \tag{1.16}$$

Physics 5C: Introductory Thermodynamics and Quantum Mechanics **Fall 2018**

Lecture 2: Statistical and Thermodynamic Concepts

Lecturer: Feng Wang *August 28* *Aditya Sengupta*

2.1 Basic Concepts

2.1.1 Probability

Previously in physics, everything was deterministic; given a force and a mass, you can predict motion for all times. Or given a charge in an electric field. This is no longer the case, because of the magnitude of the system. It's not practical to measure all 10^{23} particles in a system and determinstically find what happens to each of them. Instead, we can use the probability of microscopic states. From there, we can use statistical methods to accurately predict aggregate/average properties such as temperature and pressure, based on the probability distributions of microscopic states (their velocity, energy, etc.) as well as fluctuations around the average.

2.1.1.1 Discrete Probability Distributions

These are used when there is a discrete number of outcomes, such as rolling a die. In this case, we can separately write each of $P(1)$ (probability of the first event, e.g. rolling a one), $P(2), \ldots, P(i)$. Based on this, we can derive some general quantities and properties:

1. $\sum_i P(i) = 1$
2. The mean (or average, or expected value) $\langle x \rangle = \sum_i x_i P(i)$.

   Similarly, we can say $\langle x^2 \rangle = \sum_i x_i^2 P(i)$, or more generally $\langle f(x) \rangle = \sum_i f(x_i) P(i)$.

   For example, in a distribution defined as follows:

x	P(x)
0	$\frac{1}{2}$
1	$\frac{1}{4}$
2	$\frac{1}{4}$

   we can find an average value for x and for x^2:

$$\langle x \rangle = 0 \times 1/2 + 1 \times 1/4 + 2 \times 1/4 = 3/4 \tag{2.1}$$

$$\langle x^2 \rangle = 0^2 \times 1/2 + 1^2 \times 1/4 + 2^2 \times 1/4 = 5/4 \tag{2.2}$$

2.1.1.2 Continuous Probability Distributions

We use this when we have to define probability in an interval. This is for events that do not have a set/discrete number of potential outcomes.

Let the interval be $(x, x + \Delta x)$. Then the probability of an event in this interval is $P(x) \times \Delta x$, which in the limit as $\Delta x \to 0$ becomes $P(x)dx$. We know that the total probability is still 1; therefore if we add all the possible outcomes' probabilities, we will get 1:

$$\lim_{\Delta x \to 0} \sum_i P(x_i)\Delta x = 1 \implies \int P(x)dx = 1 \tag{2.3}$$

We can define a mean in a continuous distribution:

$$\langle x \rangle = \sum_i x_i P(x_i)\Delta x_i = \int xP(x)dx \tag{2.4}$$

Similarly, a mean square value can be found using $\int x^2 P(x)dx$, and the mean value for a function using $\int f(x)P(x)dx$.

Continuous probability also introduces the concept of variance, a measure of the fluctuation or the spread in the value. It may seem difficult to characterize this using only one quantity. To start with, we can see that it must depend upon $x - \langle x \rangle$, to see how far away from the mean a value can fall. One idea might be to take the average of this:

$$\langle x - \langle x \rangle \rangle = 0 \tag{2.5}$$

By definition of the mean, this becomes 0. We want a similar quantity that does not go negative, so that deviations from the mean only add to the variance. We could try this:

$$\langle |x - \langle x \rangle| \rangle \tag{2.6}$$

However, this is not an analytic function, so in general it is difficult to deal with. A better solution is to take a square:

$$\langle (x - \langle x \rangle)^2 \rangle = \sigma_x^2 \tag{2.7}$$

This is the *variance* of a population. To bring the variance into the same units as x, we can take a square root to get the *standard deviation*:

$$\sigma_x = \sqrt{\langle (x - \langle x \rangle)^2 \rangle} \tag{2.8}$$

This is the rms (root mean square) deviation from the mean.

2.1.2 Operations on continuous probability distributions

To characterize transformations such as finding the average (rms) momentum from a known velocity, we introduce linear transformations on distributions. Let $y = ax + b$ be a transformation on the continuous variable x. Then,

$$\langle y \rangle = \langle ax + b \rangle = a\langle x \rangle + b \tag{2.9}$$

The standard deviation after a linear transformation can also be calculated:

$$\sigma_y^2 = \langle (y - \langle y \rangle)^2 \rangle = \langle y^2 - 2y\langle y \rangle + \langle y \rangle^2 \rangle \tag{2.10}$$

$$\sigma_y^2 = \langle y^2 \rangle - 2\langle y \rangle\langle y \rangle + \langle y \rangle^2 \tag{2.11}$$

$$\sigma_y^2 = \langle y^2 \rangle - \langle y \rangle^2 \tag{2.12}$$

We can calculate these terms in x separately:

$$\langle y^2 \rangle = \left\langle (ax + b)^2 \right\rangle = \left\langle a^2x^2 + 2abx + b^2 \right\rangle \tag{2.13}$$

$$\left\langle y^2 \right\rangle = a^2 \left\langle x^2 \right\rangle + 2ab \left\langle x \right\rangle + b^2 \tag{2.14}$$

and

$$\left\langle y \right\rangle^2 = \left(a \left\langle x \right\rangle + b\right)^2 = a^2 \left\langle x \right\rangle^2 + 2ab \left\langle x \right\rangle + b^2 \tag{2.15}$$

Therefore, we can find σ_y^2:

$$\sigma_y^2 = a^2 \left(\left\langle x^2 \right\rangle - \left\langle x \right\rangle^2 \right) = a^2\sigma_x^2 \tag{2.16}$$

We see that a constant shift of b does not change the spread, but a scaling of a proportionately scales the spread.

2.1.3 Independent Variables

Suppose we have a probability distribution for velocities of a particle in the x and y directions. Consider the probability that v_x is in an interval of width dv_x, $P(v_x)dv_x$, and the same for y. If the two are independent (not correlated), the probability distribution for 2D space can be arrived at just by multiplying the two:

$$P(v_{x,y}) = P(v_x)P(v_y)dv_xdv_y \tag{2.17}$$

This is true in general for any two independent distributions (say u and v). Then, to find an overall average value for these two:

$$\langle u \cdot v\rangle = \int\int u \cdot vP_u(u)du \cdot P_v(v)dv = \int uP_u(u)du \int vP_v(v)dv\, \langle u \cdot v\rangle = \langle u\rangle \cdot \langle v\rangle \tag{2.18}$$

Similarly, $\langle u + v\rangle = \langle u\rangle + \langle v\rangle$. (This is true regardless of dependence.)

We can combine many continuous variables into one distribution. Let $y = \sum_{i=1}^{n} x_i$. Then we can calculate a mean and standard deviation of y:

$$\langle y\rangle = \langle x_1 + x_2 + \cdots + x_n\rangle = \sum_i \langle x_i\rangle = n\,\langle x\rangle \tag{2.19}$$

and

$$\sigma_y^2 = \left\langle y^2\right\rangle - \langle y\rangle^2 \tag{2.20}$$

$$\left\langle y^2\right\rangle = \left\langle (x_1 + x_2 + \cdots + x_n)^2\right\rangle = \left\langle x_1^2 + x_2^2 + \cdots + x_n^2 + x_1x_2 + x_2x_1 + \ldots\right\rangle \tag{2.21}$$

$$\left\langle y^2\right\rangle = \left\langle x_1^2\right\rangle + \left\langle x_2^2\right\rangle + \cdots + \left\langle x_n^2\right\rangle + \langle x_1\rangle\langle x_2\rangle + \langle x_2\rangle\langle x_1\rangle + \ldots \tag{2.22}$$

$$\langle y\rangle^2 = \langle x_1\rangle^2 + \langle x_2\rangle^2 + \cdots + \langle x_n\rangle^2 + \langle x_1\rangle\langle x_2\rangle + \langle x_2\rangle\langle x_1\rangle + \ldots \tag{2.23}$$

The cross terms cancel, and each x_i becomes the variance of that variable:

$$\sigma_y^2 = \left\langle y^2\right\rangle - \langle y\rangle^2 = \sigma_{x_1}^2 + \sigma_{x_2}^2 + \cdots = n\sigma_x^2 \tag{2.24}$$

2.1.4 Error Analysis

Let each measurement we take in an experiment be its own independent variable x_i. Then, define $\widetilde{x} = \frac{y}{n}$. $\langle\widetilde{x}\rangle = \langle x\rangle$, and for the standard deviation:

$$\sigma_{\widetilde{x}} = \sigma_{y/n} = \frac{1}{n}\sigma_y \tag{2.25}$$

$$\sigma_{\widetilde{x}} = \frac{1}{n}\sqrt{\sigma_y^2} = \frac{\sqrt{n\sigma_x^2}}{n} = \frac{\sigma_x}{\sqrt{n}} \tag{2.26}$$

2.2 Random walk

Consider a number line, on which a person is walking randomly. There is an equal chance that the person takes a step forward (+1) or backward (-1) at each step. What is the expected value of the person's position after N steps?

We see that $\langle x \rangle = 0$ and $\langle x^2 \rangle = 1 = \sigma_x^2$. After N steps, $\langle y \rangle = 0$ (the sum of N independent random steps) and $\sigma_y = \sqrt{N}$.

This is a central idea in understanding concepts related to thermal physics such as Brownian motion.

2.3 Heat and Temperature

Heat is related to a transfer of energy. More specifically, it is defined as energy in transit, usually from a hot subject to a cold object. The unit of heat is the same as that of energy, joules (J), and the rate of heating is the same units as power, watts (W).

We define the heat capacity as the amount of heat needed to raise the temperature of an object by one unit of temperature.

$$C = \frac{dQ}{dT} \tag{2.27}$$

The specific heat (c) is defined as the heat capacity per unit mass.

Systems can be set up to have constant volume or constant pressure, as the nature of how the system changes is different in both of these cases. Therefore, we define two heat capacities:

$$C_v = \frac{\partial Q}{\partial T}_v \tag{2.28}$$

$$C_p = \frac{\partial Q}{\partial T}_p \tag{2.29}$$

Usually $C_p > C_v$, as we need more heat to raise temperature under constant pressure.

Physics 5C: Introductory Thermodynamics and Quantum Mechanics **Fall 2018**

Lecture 3: Temperature

Lecturer: Feng Wang *30 August* *Aditya Sengupta*

3.1 Thermal Equilibrium

Intuitively, we understand that a system comes to thermal equilibrium if a hot object of temperature T_1 is connected to a cool object of temperature T_2, and over time t they equalize to both attain temperature T_f. At this point, there is no net flow of heat between the objects. Here, an irreversible process has taken place. It intuitively seems like this process should be reversible, as it is only a transfer of energy, but this is not the case (probabilistically).

This idea of gaining equilibrium also leads to the 0th law of thermodynamics: if a system A is in equilibrium with B, and A is in equilibrium with some other system C, then B and C are in equilibrium too.

(This is Prof. Norman Yao)

In principle, a quantum mechanical description of a thermodynamic system *would* allow for oscillation between the two states (T_1, T_2) and (T_f, T_f), but the timescale for oscillation is astronomically large even by astronomically large standards.

The 0th law helps us to measure temperature exponentially. We can place a body in an unknown T, and use a calibration system that has a measurable property that exhibits a well-known dependence on temperature. This is the concept behind a thermometer. Note that a thermometer should have a heat capacity much smaller than that of the heat being measured. Recall that heat capacity is the heat required to cause a unit change in temperature, $C = \frac{dQ}{dT}$. This is so that attaining thermal equilibrium does not appreciably change the energy of the system by measuring it.

The ideal gas law, $PV = nRT$, can also be used to measure temperature. Either pressure is held fixed and changes in volume are measured, or vice versa.

All of the ways in which we are defining temperature depend on some measurable property. In general, almost all measurable properties are non-linear. For example, the direct pressure/volume relationship does not hold when the liquid being studied freezes. Defining temperature in terms of physically measurable properties would therefore be painful. Instead, we want to find a theoretical - in this case statistical - basis for defining temperature.

3.2 Micro/Macro States

Consider a box containing 100 fair coins. We shake the box and look at the configuration. There are 2^{100} possible configurations, which is about 10^{30}. Each configuration (set of HTTHH...TH) is a microstate, but that probably isn't how you would characterize the system; you might say something more like “there are 53 heads and 47 tails”. This is a macrostate.

The number of microstates corresponding to a macrostate can be very large. In this example, there are ${}^{100}C_{50} = \frac{100!}{50!50!} \approx 4 \times 10^{27}$ microstates for the one macrostate 50H, 50T. It can also be small; the macrostate 100H, 0T has only one microstate.

We can describe or measure a system either in terms of many equally likely microstates, or a smaller number of macrostates that are not all equally likely. The most likely macrostate for a system is the one with the most configurations or microstates.

A thermal analogy for this would be to specify a microstate for a thermal system (the position/velocity of each atom), which is operationally impossible for any realistic numbe of atoms in a gas. Instead, we generally describe a thermal system via its macrostate. There are several possible ways of going about this, such as total energy, or volume and pressure. To build a statistical notion of temperature, we consider a system of energy E and imagine there exist $\Omega(E)$ microstates. This allows us to return to the two systems in thermal contact coming into equilibrium, meaning that some notion of temperature has become the same between the two. Let there be $\Omega_1(E_1)$ microstates of the first system, and $\Omega_2(E_2)$ microstates of the second. For a total energy of $E_1 + E_2$, the total number of microstates is $\Omega_1(E_1) \times \Omega_2(E_2)$.

Systems can exchange energy and eventually come into equilibrium, in a process called *thermalization*. The key insight here is that a system will equilibrate to a microstate such that the number of microstates $\Omega_1(E_1) \times \Omega_2(E_2)$ is maximized.

In this model, we assume that each microstate is equally likely, and the dynamics of a system allows us to constantly move between microstates. Additionally, given enough time for the system to equilibrate, the system will explore all its microstates.

These assumptions only imply that a system will most likely be found in a configuration with the largest number of microstates.

3.3 Definition of temperature

We want to maximize $\Omega_1(E_1) \times \Omega_2(E_2)$. Without loss of generality, we choose to shift E_1. Then, we know that at the maximum energy,

$$\frac{\partial}{\partial E_1}\left(\Omega_1(E_1) \times \Omega_2(E_2)\right) = 0 \tag{3.1}$$

We apply the product rule and chain rule:

$$\Omega_2(E_2)\frac{\partial \Omega_1(E_1)}{\partial E_1} + \Omega_1(E_1)\frac{\partial \Omega_2(E_2)}{\partial E_2}\frac{\partial E_2}{\partial E_1} = 0 \tag{3.2}$$

We can also use the fact that $E_1 + E_2 = E$, a constant total energy. Taking a derivative tells us that $dE_1 = -dE_2 \implies \frac{dE_2}{dE_1} = -1$, therefore:

$$\frac{1}{\Omega_1}\frac{\partial \Omega_1}{\partial E_1} - \frac{1}{\Omega_2}\frac{\Omega_2}{E_2} = 0 \tag{3.3}$$

This can be expressed as the derivative of natural log:

$$\frac{d \ln \Omega_1}{dE_1} = \frac{d \ln \Omega_2}{dE_2} \tag{3.4}$$

This serves as a fundamental definition for temperature:

$$\frac{1}{k_B T} = \frac{d \ln \Omega}{dE} \tag{3.5}$$

and the above equation is the condition for thermal equilibrium. This also leads to a fundamental definition for entropy, S:

$$S = k_B \ln \Omega \frac{1}{T} = \frac{\partial S}{\partial E} \tag{3.6}$$

k_B, a proportionality constant, is $1.38 \times 10^{-23} \frac{J}{K}$.

3.4 Ensembles

The notion of ensembles is introduced to deal with the number of possible states of a system there are; either we repeat an experiment many times, or consider many copies of an identical system all described by the same temperature. There are three ensembles, of which we will only really work with one.

3.4.1 Microcanonical Ensemble

This is an ensemble of systems at the same energy.

3.4.2 Canonical Ensemble

This is an ensemble of systems, each of which can exchange energy with a large bath.

3.4.3 Grand Canonical Ensemble

This is an ensemble of systems, each of which can exchange both energy and particles with a large bath. The exchange of particles allows us to define a *chemical potential* later in the course.

3.5 The Canonical Ensemble

Imagine we have two systems coupled, connected to a bath at temperature T. The whole system has energy E, of which the system has energy ε and the bath has energy $E - \varepsilon$. We assume that the bath is unaffected by the system. The bath has $\Omega(E - \varepsilon)$ microstates.

We want to know the probability that the system has energy ε. This is proportional to the number of microstates of the bath (why?).

We can apply the above relation:

$$\frac{1}{k_BT} = \frac{d\ln\Omega(E-\varepsilon)}{dE} \tag{3.7}$$

Then, we can Taylor expand:

$$\ln\Omega(E-\varepsilon) = \ln\Omega(E) - \frac{d\ln\Omega(E)}{dE}\varepsilon + \dots \tag{3.8}$$

$$\ln\Omega(E-\varepsilon) = \ln\Omega - \frac{\varepsilon}{k_BT} \tag{3.9}$$

$$\therefore \Omega(E-\varepsilon) = \Omega(E)e^{-\varepsilon/k_BT} \tag{3.10}$$

$$\therefore P(\varepsilon)\ e^{-\varepsilon/k_BT} \tag{3.11}$$

The exponential above is known as the Boltzmann factor, and this relation is known as a canonical or a Boltzmann distribution.

We can normalize this distribution:

$$P(stater) = \frac{e^{-E_r/k_BT}}{\sum_i e^{-E_i/k_BT}} \tag{3.12}$$

The denominator here is commonly denoted with Z, and is often known as the partition function.

Example

(Note that $\beta = \frac{1}{k_BT}$, which lets us represent the Boltzmann factor as $e^{-\beta\varepsilon}$).

Let there be two states: $\varepsilon = 0$ and $\varepsilon > 0$. The average energy of the system is a probability-weighted average:

$$\langle E\rangle = \sum_i \varepsilon_i P_i \tag{3.13}$$

So, we calculate the probabilities:

$$P_{\varepsilon=0} = \frac{1}{1+e^{-\beta\varepsilon}} \tag{3.14}$$

$$P_{\varepsilon>0} = \frac{e^{-\beta\varepsilon}}{1+e^{-\beta\varepsilon}} \tag{3.15}$$

Therefore, the average energy is

$$\langle E\rangle = \frac{e^{-\beta\varepsilon}}{1+e^{-\beta\varepsilon}} \tag{3.16}$$

$$\langle E\rangle = \frac{\varepsilon}{1+e^{\beta\varepsilon}} \tag{3.17}$$

At high temperatures, $\frac{\varepsilon}{kT}$ is low, therefore (finish up)

Physics 5C: Introductory Thermodynamics and Quantum Mechanics **Fall 2018**

Lecture 4: Temperature, Laws of Thermodynamics

Lecturer: Feng Wang *September 4* *Aditya Sengupta*

We saw last time that a macrostate could be made up of many microstate configurations. Since each microstate has equal probability, thermal equilibrium is reached at the macrostate with the largest number of microstates.

We can characterize a macrostate in terms of $\Omega(V, T, E)$; this function is the number of microstates. The configuration with the maximum Ω gives thermal equilibrium. Using this concept, we previously derived this relation between a number of microstates at thermal equilibrium and temperature:

$$\frac{d\ln\Omega}{dE} = k = \frac{1}{k_B T} \tag{4.1}$$

k_B has units of energy/temperature, or $\frac{J}{K}$, and the numerical value is 1.38×10^{-23}.

Based on the concept of the canonical ensemble, in which a system is connected to a heat reservoir, we can find the probability of each microstate:

$$P(microstate) = \frac{e^{-\frac{E}{k_B T}}}{\sum_i e^{-\frac{E_i}{k_B T}}} \tag{4.2}$$

The numerator is called the Boltzmann factor, and the denominator is Z, the partition function. The whole expression defines the *Boltzmann distribution.*

4.1 Examples of the Boltzmann distribution

4.1.1 Atmosphere

Consider an isothermal (equal temperature) atmosphere, with a number of molecules h. We want to find the distribution of air density in this atmosphere. This can be done with classical mechanics, but the Boltzmann distribution presents an easier way.

At a height z, the (potential) energy is $E = mgz$. Therefore

$$P(z) \propto e^{-\frac{mgz}{k_B T}} \tag{4.3}$$

This is the probability of finding a molecule at height z. We can make a similar function for density that is proportional to this probability,

$$n(z) = Ae^{-\frac{mgz}{k_B T}} n(0) = A \implies n(z) = n(0)e^{-\frac{mgz}{k_B T}} \tag{4.4}$$

4.1.2 Chemical Reactions

For H_2 and O_2 to combine to form H_2O, there is an energy barrier to be crossed, on the order of 0.5 eV. The k_BT at room temperature is 25×10^{-3} eV, therefore thermal energy usually needs to be added to the system to cross the barrier. However, there is a very small probability that the system will naturally attain that energy, which is described by the Boltzmann distribution. The probability is proportional to

$$r_0 = e^{-\frac{E_{act}}{k_B T}} \tag{4.5}$$

Consider a fridge that lowers the temperature of its contents by ΔT. Then, the activation energy becomes

$$r' = e^{-\frac{E_{act}}{k_B(T-\Delta T)}} \tag{4.6}$$

We can rearrange and Taylor expand the denominator of r' to get:

$$r' = e^{-\frac{E_a}{k_B T}\left(1+\frac{\Delta T}{T}\right)} \frac{r'}{r_0} = e^{-\frac{E_a}{k_B T}\frac{\Delta T}{T}} \tag{4.7}$$

With some reasonable assumptions, we get $r_0 \approx 8r'$, meaning reactions (e.g. food being spoiled) take place 8 times more slowly in a fridge. If something would have spoiled in 1 day outside, it will spoil in 8 days inside.

4.2 State Variables

Let a system be anything we want to study, and let its surroundings be everything else (treated as a thermal reservoir). We will consider the state of thermal equilibrium for now, which we can characterize through macroscopic observable vaues. Functions of state (that do not depend on the system's history) include temperature, pressure, energy, and volume; these are ways of indirectly measuring heat or work.

Consider a system consisting of state variables $x_n = (x_1, x_2, \cdots)$. Let $f(x)$ be a function of state. Then, if a state variable changes $x_i \to x_f$, then

$$\Delta f = \int_{x_i}^{x_f} df = f(x_f) - f(x_i) \tag{4.8}$$

For example, let there be two state variables x and y that completely define a system, and let $f(x, y) = xy$. Then, any path between $(0, 0)$ and $(1, 1)$ wil cause this change in f:

$$\Delta f = \int_{(0,0)}^{(1,1)} df = [xy]_{(0,0)}^{(1,1)} = 1 \tag{4.9}$$

Consider another variable defined in terms of a differential, $dg = ydx$. Here, the path can be a factor. Integrate between the same two points along the line $y = x$:

$$\int_{(0,0)}^{(1,1)} ydx = \int_0^1 xdx = \frac{1}{2} \tag{4.10}$$

whereas along $y = 0$ followed by $x = 1$:

$$\int_{x=0}^{x=1} ydx + \int_{y=0}^{y=1} ydx = 0 + 0 = 0 \tag{4.11}$$

These cannot be used to define state variables, as they are not conservative.

4.3 First Law of Thermodynamics

We want to find a relation between work/heat and intrinsic state variables. The law provides this; it states that **heat and work are both forms of energy**, or that **total energy is conserved**.

We can define an internal energy U, the sum of all the energies in the system. Then, we can say that $\Delta U = \Delta Q + \Delta W$: any change in energy must come only from the changes in heat or work.

Consider a thermally isolated system in which $\Delta Q = 0$, so $dU = dW$ (the d in dW should have a bar across it to indicate W isn't a conservative state variable, but eh), and we compress the system, changing the volume and therefore the pressure. We know $dW = \vec{F} \cdot d\vec{x}$, and therefore

$$dW = -\frac{F}{A}(Adx) = -\frac{PA}{A}dV = -PdV \tag{4.12}$$

(with a negative because volume is decreasing). Assuming no friction and everything in equilibrium, this relation holds.

4.4 Heat capacity of a gas system

Consider a gas system in which energy is dependent on temperature and volume: $U = U(T, V)$. Then

$$dU = \left(\frac{\partial U}{\partial T}\right)_V dT = \left(\frac{\partial U}{\partial V}\right)_T dV \tag{4.13}$$

Then, a change in the system from $U(T_1, V_1)$ to $U(T_2, V_2)$ causes a change in energy ΔU, which can be described by the above equation in terms of the changes in T and U.

To find $\frac{dQ}{dT}$, we can use the First Law and the equation for dW derived above:

$$dQ = dU + PdV = \left(\frac{\partial U}{\partial T}\right)_V dT + \left[\left(\frac{\partial U}{\partial V}\right)_T + P\right] dV \tag{4.14}$$

We can simplify these expressions by defining constants:

$$C_v = \left(\frac{\partial Q}{\partial T}\right)_V = \left(\frac{\partial U}{\partial T}\right)_V \tag{4.15}$$

$$C_p = \left(\frac{\partial U}{\partial T}\right)_V + \left[\left(\frac{\partial U}{\partial T}\right)_V + P\right]\left(\frac{\partial V}{\partial T}\right)_P \tag{4.16}$$

$$C_p - C_v = \left[\left(\frac{\partial U}{\partial V}\right)_T + P\right]\left(\frac{\partial V}{\partial T}\right)_P \tag{4.17}$$

C_p is usually larger than C_v.

In an ideal gas, $U(V,T) = U(T) = \frac{3}{2}RT$. The $\frac{3}{2}$ is because molecules usually have three degrees of freedom, and the volume independence comes from the Boltzmann expression, that only has T dependence and not V dependence. Then,

$$C_v = \left(\frac{\partial U}{\partial T}\right)_V = \frac{3}{2}R \tag{4.18}$$

$$C_p - C_v = \left[\left(\frac{\partial U}{\partial V}\right)_T + P\right]\left(\frac{dV}{dT}\right)_P = \left(\frac{PdV}{dT}\right)_P = R \tag{4.19}$$

$$C_p = C_v + R = \frac{5}{2}R \tag{4.20}$$

We then define $\gamma = \frac{C_p}{C_v} = \frac{5}{3}$, the adiabatic index. It is usually > 1.

Physics 5C: Introductory Thermodynamics and Quantum Mechanics **Fall 2018**

Lecture 5: Thermal Processes

Lecturer: Feng Wang *6 September* *Aditya Sengupta*

5.1 Reversible Isothermal Processes and Adiabatic Processes

5.1.1 Reversible Isothermal Processes

The most straightforward way to have a reversible process is to move the system very slowly, so that it essentially remains in equilibrium throughout the process. This is referred to as a *quasi-static* process.

In a reversible isothermal process, the change in temperature is zero: $\Delta T = 0$. (We can do this by placing the system in a thermal reservoir.) For an ideal gas, in which $U = \frac{3}{2}nRT$, $dU = 0$ as well. Therefore $dW = -dQ$, meaning we can quantify the change in heat over time:

$$\Delta Q = \int_{state1}^{state2} dQ = \int_1^2 -dW = \int_1^2 PdV \tag{5.1}$$

from which we can use the ideal gas law:

$$\Delta Q = \int_1^2 \frac{RT}{V} dV = RT \ln \frac{V_2}{V_1} \tag{5.2}$$

5.1.2 Adiabatic Expansion

An adiabatic system has no heat exchange with its surroundings, which can be achieved by isolating the system. Therefore $\Delta Q = 0$, which means

$$dU = dW \tag{5.3}$$

$$C_v dT = -PdV \tag{5.4}$$

$$C_v \frac{dT}{T} = -R\frac{dV}{V} \tag{5.5}$$

$$C_v \int \frac{dT}{T} = -R \int \frac{dV}{V} \tag{5.6}$$

Let the initial state be characterized by V_1, T_1 and the final state by V_2, T_2. Then we can set bounds on the integrals:

$$C_v \ln T \Big|_{T_1}^{T_2} = -R \ln V \Big|_{V_1}^{V^2} \tag{5.7}$$

$$\ln \frac{T_2}{T_1} = -\frac{R}{C_v} \ln \frac{V_2}{V_1} \tag{5.8}$$

$$R = C_p - C_v \implies \frac{R}{C_v} = \frac{C_p}{C_v} - 1 = \gamma - 1 \tag{5.9}$$

$$\therefore \ln \frac{T_2}{T_1} = -(\gamma - 1) \ln \frac{V_2}{V_1} \tag{5.10}$$

$$\ln T_2 = (\gamma - 1) \ln V_2 = \ln T_1 + (\gamma - 1) \ln V_1 \tag{5.11}$$

$$\tag{5.12}$$

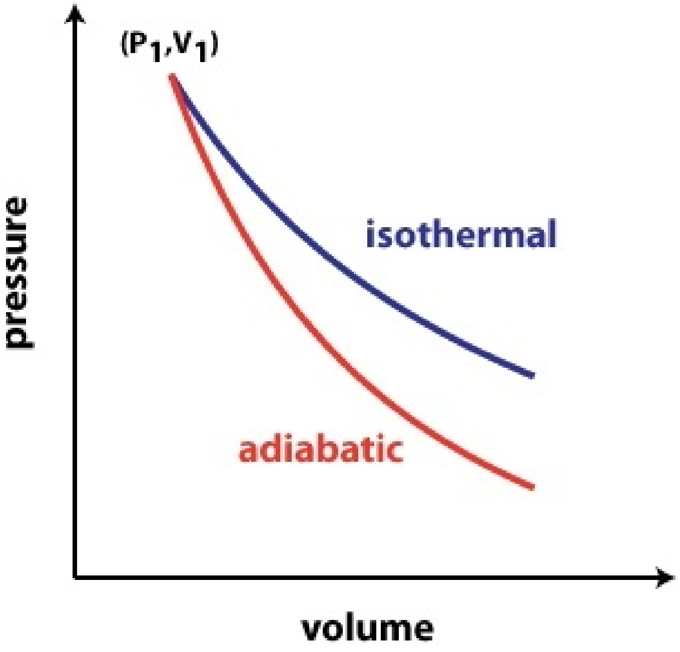
This tells us that the quantity $TV^{\gamma-1}$ is a constant for a system.

We can apply the ideal gas law again to give us a constant relation in volume and pressure:

$$PV = nRT \tag{5.13}$$

$$\frac{PV}{nR} \cdot V^{\gamma-1} = \frac{PV^{\gamma}}{nR} = \text{constant} \tag{5.14}$$

$$\therefore PV^{\gamma} = \text{constant.} \tag{5.15}$$



5.1.3 Adiabatic atmosphere

The previous assumption that the atmosphere is isothermic was not very good. The atmosphere is much closer to adiabatic.

We want to use the hydrostatic equation to find an expression for atmospheric pressure as a function of z, which can easily be found via balancing forces:

$$\Delta P A = -\left(\rho V\right) g = -\rho g A \Delta z \tag{5.16}$$

$$dP = -\rho g dz \tag{5.17}$$

$$\tag{5.18}$$

Let $\rho = \frac{N}{V} \cdot m_0$, the mass in a unit volume due to one mole multiplied by the number of moles. The ideal gas law tells us that this is

$$\rho = \frac{PV}{k_B T} \cdot m_0 \tag{5.19}$$

$$\tag{5.20}$$

Therefore, we can substitute this back into the hydrostatic equation to get

$$dP = -\frac{P}{k_B T} mg dz \tag{5.21}$$

$$T\frac{dP}{P} = -\frac{mg}{k_B} dz \tag{5.22}$$

Note that if we hold temperature constant, we find that pressure decays exponentially with z, which is the result we got in the isothermal case. We can now apply the adiabatic condition,

$$P\left(\frac{RT}{P}\right)^{\gamma} = \text{constant} \tag{5.23}$$

$$P^{1-\gamma} \cdot T^{\gamma} = \text{constant} \tag{5.24}$$

$$\tag{5.25}$$

We can take a derivative,

$$(1-\gamma)\frac{dP}{P} + \gamma\frac{dT}{T} = 0 \tag{5.26}$$

$$\frac{dP}{P} = -\frac{\gamma}{1-\gamma}\frac{dT}{T} = \frac{\gamma}{\gamma-1}\frac{dT}{T} \tag{5.27}$$

$$\tag{5.28}$$

Having isolated $\frac{dP}{P}$, we can substitute this back into the equation with a differential in z from above:

$$T\frac{\gamma}{\gamma-1}\frac{dT}{T} = -\frac{mg}{k_B}dz \tag{5.29}$$

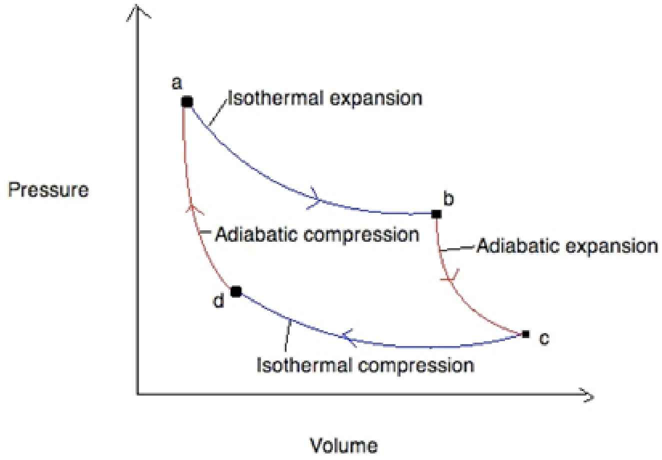
$$\frac{\gamma}{\gamma-1}dT = -\frac{mg}{k_B}dz \tag{5.30}$$

That is, $\frac{dT}{dz}$ is constant, and negative.

5.2 Heat Engines and the Second Law

The Second Law of Thermodynamics is the first concept in physics that is time-dependent. There are a couple of different ways of stating it; Clausius's statement is *no process whose sole result is to transfer heat from cold to hot is possible*, and Kelvin's is *no process whose sole result is complete conversion of heat to work is possible.*

A Carnot engine is a simplification of a real heat engine, retaining only the essential parts. This makes intuitively studying it much easier. It is a system operating a cyclic process that converts heat to work.



Because $dW = PdV$, the work done by a Carnot engine is the area enclosed in the above P/V diagram.

Let the four corners of the quadrilateral describing the Carnot engine be A through D. From A to B (at $T = T_h$) and from C to D (at $T = T_l$), an isothermal process takes place; from B to C and from D to A, it is adiabatic.

$$A \to B \implies Q_h = RT_h \ln \frac{V_B}{V_A} \tag{5.31}$$

$$B \to C \implies T_h V_B^{\gamma-1} = T_l V_C^{\gamma-1} \implies \frac{T_h}{T_l} = \left(\frac{V_C}{V_B}\right)^{\gamma-1} \tag{5.32}$$

$$C \to D \implies Q_l = -RT_l \ln \frac{V_D}{V_C} \tag{5.33}$$

$$D \to A \implies \frac{T_l}{T_h} = \left(\frac{V_A}{V_D}\right)^{\gamma-1} \tag{5.34}$$

We also know that $\Delta U = 0$ after a full cycle, therefore $\Delta Q = -\Delta W$.

We can combine the two adiabatic relations to get

$$\left(\frac{V_C}{V_B}\right)^{\gamma-1} = \left(\frac{V_D}{V_A}\right)^{\gamma-1} \tag{5.35}$$

$$\frac{V_C}{V_B} = \frac{V_D}{V_A} \implies \frac{V_B}{V_A} = \frac{V_C}{V_D} \tag{5.36}$$

and combining the isothermal relations

$$\frac{Q_h}{Q_l} = \frac{T_h}{T_l} \frac{\ln \frac{V_B}{V_A}}{\ln V_C V_D} \tag{5.37}$$

$$\frac{Q_h}{Q_l} = \frac{T_h}{T_l} \tag{5.38}$$

which gives us

$$\frac{W}{Q_h} = \frac{Q_h - Q_l}{Q_h} = 1 - \frac{T_l}{T_h} = \eta \tag{5.39}$$

η is the Carnot efficiency. This is the maximum possible efficiency a heat engine can attain.

Physics 5C: Introductory Thermodynamics and Quantum Mechanics **Fall 2018**

Lecture 6: Entropy and Heat Engines

Lecturer: Feng Wang *11 September* *Aditya Sengupta*

Carnot's theorem states that no possible engine between given temperatures T_h and T_l can have greater efficiency than a Carnot engine. This can be proved using the second law of thermodynamics. Let there be an engine E with efficiency $\eta > \eta_{Carnot}$ between T_h and T_l. Use this engine to do work, which can be given as input to the corresponding Carnot engine to move from T_l to T_h. The total work in this system is zero, and the system returns to its original state (T_h to T_l with engine E, then T_l to T_h using the Carnot engine).

We claim that $\eta > \eta_{Carnot}$, which implies

$$\frac{W}{Q_{h,E}} > \frac{W}{Q_h} \implies Q_h - Q_{h,E} > 0 \tag{6.1}$$

which implies more heat is released than absorbed. Therefore there is a transfer of heat from T_l to T_h. Either there is heat dissipation in which case $\eta < \eta_{Carnot}$, or there is no net work and no heat transfer, in which case $\eta = \eta_{Carnot}$, or heat flows so as to violate the second law.

A corollary of this is that all reversible engines must have the same efficiency as a Carnot engine. (see above)

6.1 Equivalence of Second Law Statements

We can prove that if a system violates Kelvin's statement of the law, it must violate Clausius' statement. Consider a system that converts some input heat Q_h entirely to work W, such that its efficiency its $\eta = \frac{W}{Q_h} = 1$. Therefore, if a system completely converts heat to work, its efficiency is greater than the Carnot efficiency.

We can also easily show that violating Clausius' statement would violate Kelvin's statement too. Consider a system that absorbs heat at low temperature and radiates it at high temperature without doing any work. Then the total absorbed heat is $Q_h - Q_l$. This is a system that completely converts this heat to work, therefore it violates Kelvin's statement.

6.2 Reverse Heat Engine

We want an engine that can absorb heat at low temperature and give heat at high temperature, with the minimum possible work input. (A fridge)

$$\eta_R = \frac{Q_L}{W} = \frac{Q_l}{Q_h - Q_l} = \frac{T_l}{T_h - T_l} \tag{6.2}$$

Similarly, we can look at a heat pump, which is similar but it generates heat to (for example) put into a room.

$$\eta_R = \frac{Q_h}{W} = \frac{Q_h}{Q_h - Q_l} = \frac{T_h}{T_h - T_l} \tag{6.3}$$

6.3 Clausius's theorem

In a Carnot engine, we know that

$$\begin{align} \frac{Q_h}{Q_l} &= \frac{T_h}{T_l} \tag{6.4} \\ \frac{Q_h}{T_h} - \frac{Q_l}{T_l} &= 0 \tag{6.5} \\ \sum_{cycle} \frac{\Delta Q_i}{T_i} = \frac{Q_h}{T_h} + \frac{(-Q_l)}{T_l} &= 0 \tag{6.6} \\ \oint \frac{dQ}{T} &= 0 \tag{6.7} \end{align}$$

The above is true for a Carnot cyle. Consider a large number of Carnot engines placed one after another, each making an infinitesimal change in temperature to produce a net ΔW consisting of many dW components. Let the initial working temperature be T, then the change in heat at step i is

$$\frac{dQ_i}{T_i} = \frac{dQ_{h,i}}{T} = \frac{dQ_i + dW_i}{T} \implies dW_i = dQ_i \left(\frac{T}{T_i} - 1 \right) \tag{6.8}$$

Then, by the second law, our total work is ≤ 0, so

$$W_{total} = \sum_i dW_i + \Delta W \leq 0 \tag{6.9}$$

By the first law,

$$\Delta W = \sum_i dQ_i \tag{6.10}$$

Therefore

$$\sum (dW_i + dQ_i) < 0 \sum \left[\left(\frac{T}{T_i} - 1 \right) dQ_i + dQ_i \right] T \sum_i \frac{dQ_i}{T_i} = T \oint \frac{dQ_i}{T_i} \leq 0 \tag{6.11}$$

6.4 Entropy

$\oint \frac{dQ_{rev}}{T} = 0 \implies \int_A^B \frac{dQ_{rev}}{T}$ is path independent. Therefore it is an exact differential. So, we introduce a new function of state S:

$$dS = \frac{dQ_{rev}}{T} \tag{6.12}$$

$$\int_A^B \frac{dQ_{rev}}{T} = S(B) - S(A) \tag{6.13}$$

6.5 Irreversible Processes

An irreversible process has the characteristic

$$\oint \frac{dQ}{T} \leq 0 \tag{6.14}$$

Therefore

$$\int_A^B \frac{dQ}{T} + \int_B^A \frac{dQ_{rev}}{T} \leq 0 \tag{6.15}$$

$$\int_A^B \frac{dQ}{T} \leq -\int_B^A \frac{dQ_{rev}}{T} = \int_A^B \frac{dQ_{rev}}{T} \tag{6.16}$$

$$0 = \int_A^B \frac{dQ}{T}(\text{isolated system}) \leq \int_A^B \frac{dQ_{rev}}{T} = \int_A^B dS = S(B) - S(A) \tag{6.17}$$

We can now revisit the first law to find:

$$dU = dQ + dW \tag{6.18}$$

$$dU = TdS - PdV \tag{6.19}$$

$$\frac{1}{k_B T} = \frac{d\ln\Omega}{dU} \implies \frac{1}{k_B}\left(\frac{\partial S}{\partial U}\right)_V \tag{6.20}$$

which gives us something about microstates. (Revisit)

Physics 5C: Introductory Thermodynamics and Quantum Mechanics **Fall 2018**

Lecture 8: Micro and Macro Definitions of Entropy

Lecturer: Feng Wang *September 18* *Aditya Sengupta*

8.1 Mathematical Relations between Thermodynamic Quantities

We've seen the first law of thermodynamics many times by now:

$$dU = dQ + dW \tag{8.1}$$

In a reversible process, we have expressions for dQ and dW:

$$dU = TdS - PdV \tag{8.2}$$

Since these are all exact differentials, this expression holds true even for irreversible processes. This lets us express U as a function of entropy and volume $U(S, V)$. This lets us rewrite dU by differentiating:

$$dU = \left(\frac{\partial U}{\partial S}\right)_V dS + \left(\frac{\partial U}{\partial V}\right)_S dV \tag{8.3}$$

Comparing coefficients tells us that

$$\left(\frac{\partial U}{\partial S}\right)_V = T \tag{8.4}$$

$$\left(\frac{\partial U}{\partial V}\right)_S = -P \tag{8.5}$$

$$\tag{8.6}$$

and therefore

$$\frac{P}{T} = -\left(\frac{\partial U}{\partial V}\right)_S \left(\frac{\partial S}{\partial U}\right)_V \tag{8.7}$$

Also, we can rewrite the original equation coming from the First Law,

$$dS = \frac{dU}{T} + \frac{P}{T}dV \tag{8.8}$$

and therefore comparing coefficients here tells us

$$\left(\frac{\partial S}{\partial U}\right)_V = \frac{1}{T} \tag{8.9}$$

$$\left(\frac{\partial S}{\partial V}\right)_U = \frac{P}{T} \tag{8.10}$$

8.2 Joule expansion

This is the term for the free expansion of gas. Consider two sides of a container, each of volume V_i separated by a valve, in which an ideal gas exerting pressure P_i is confined to the left side. The valve is opened and the gas is allowed to expand. No work is done on the system, and no heat is exchanged with the outside. Therefore $\Delta U = 0$ by the first law. We also know that $U = \frac{3}{2}RT$, so this process is isothermal; $T_f = T_i$. We also know from the ideal gas law that

$$P_i V_i = RT_i \tag{8.11}$$

$$P_f(2V_i) = RT_i \tag{8.12}$$

$$\therefore P_f = \frac{P_i}{2} \tag{8.13}$$

This analysis does not tell us about the change in entropy. But since entropy is path-independent, we can find the change in entropy due to any reversible process that takes the system from (V_i, P_i, T_i) to $(2V_i, \frac{P_i}{2}, T_i)$. For isothermal expansion, we know that

$$dS = \frac{dU}{T} + \frac{PdV}{T} \tag{8.14}$$

Since the process is isothermal, $dU = 0$, therefore by the ideal gas law we can reduce this to

$$dS = R\frac{dV}{V} \tag{8.15}$$

Integrating:

$$\Delta S = \int_i^f dS = \int_{V_i}^{2V_i} R\frac{dV}{V} = R\ln\frac{2V_i}{V_i} = R\ln 2 \tag{8.16}$$

This connects to the $k_B \ln \Omega$ definition of entropy as well; in a system with N_A particles, there are 2^{N_A} microstates (every particle is either in the left or the right partition), therefore change in entropy based on that definition is

$$S_i = k_B \ln 1 = 0 \tag{8.17}$$
$$S_f = k_B \ln 2^{N_A} = N_A k_B \ln 2 = R \ln 2 \tag{8.18}$$
$$\Delta S = S_f - S_i = R \ln 2 \tag{8.19}$$

In a system in which there are n_0 starting states and $2n_0$ ending states, this becomes

$$S_i = k_B \ln n_0^{N_A} = R \ln n_0 \tag{8.20}$$
$$S_f = k_B \ln (2n_0)^{N_A} = R \ln 2n_0 \tag{8.21}$$
$$\Delta S = S_f - S_i = R \ln \frac{2n_0}{n_0} = R \ln 2 \tag{8.22}$$

8.3 Gas Mixing

Let a system with two partitions both at (P, T) and with volumes χV_o and $(1-\chi)V_o$ be connected by a closed valve. What is the change of entropy when the valve is opened?

From the ideal gas law, we know that the number of moles in each side must be χN_o and $(1-\chi)N_o$ respectively. This lets us set up a similar integral over volume for the entropy change in isothermal expansion:

$$\Delta S = \int_{\chi V_o}^{V_o} (\chi N_o k_B) \frac{dV}{V} = \int_{(1-\chi)V_o}^{V_o} (1-\chi) N_o k_B \frac{dV}{V} \tag{8.23}$$
$$\Delta S = -N_o k_B \left(\chi \ln \chi + (1-\chi) \ln (1-\chi) \right) \tag{8.24}$$

Since $\chi < 1$, the negative sign outside will be cancelled. In the case $\chi = \frac{1}{2}$, $\Delta S = N_o k_B \ln 2$ which is similar to the previous expressions.

8.4 Entropy

Consider two systems, one at (P_1, T_1) and the other at (P_2, T_2), that can exchange heat and volume. We want to find the equilibrium conditions of the two systems based on entropy.

We know

$$dS = \frac{dU}{T} + \frac{P}{T} dV \tag{8.25}$$

and we can set up an expression for the total change in entropy:

$$\Delta S_{total} = \Delta S_1 + \Delta S_2 = \frac{\Delta U}{T_1} + \frac{P_1 \Delta V}{T_1} + \frac{-\Delta U}{T_2} + \frac{P_2}{V_2}(-\Delta V) \tag{8.26}$$
$$\Delta S_{total} = \Delta U \left(\frac{1}{T_1} - \frac{1}{T_2} \right) + \Delta V \left(\frac{P_1}{T_1} - \frac{P_2}{T_2} \right) \tag{8.27}$$

At thermal equilibrium, $\Delta S = 0$.

8.5 Entropy and probability

Given a two-part system that we can characterize by $\langle U\rangle, \langle S\rangle, \langle P\rangle$, we want to find the entropy in part 1 and 2 ($\langle S\rangle$).

$$S = k_B \ln \Omega \tag{8.28}$$

$$\langle S_2\rangle = \sum_i P_i S_i = \sum_i P_i k_B \ln n_i \tag{8.29}$$

Let the whole system $(1+2)$ have N microstates, and let (2) have n_i microstates. Then the probability of a particle being in 2 is $P_i = \frac{n_i}{N}$. This allows us to define an average entropy for part 1:

$$\langle S_1\rangle = S_{total} - S_2 = k_B(\sum_i P_i)\ln N - \sum_i P_i k_B \ln n_i = k_B \sum_i P_i \ln \frac{1}{P_i} \tag{8.30}$$

Physics 5C: Introductory Thermodynamics and Quantum Mechanics **Fall 2018**

Lecture 9: Democracy doesn't work (microstate distributions, equipartition)

Lecturer: Feng Wang *September 20* *Aditya Sengupta*

We saw previously that

$$S = k_B \ln \Omega \tag{9.1}$$

However, this assumes that every microstate has equal probability, which is not the case. This definition of entropy was built on a binomial distribution of microstates, assuming each one was equally likely (analogous to flipping a coin, etc.), but that isn't always true. For example, the distribution of particles in a box with a gravitational field attracting it to one side is not binomial.

We saw previously that the entropy of a system can be expressed in terms of the probability of microstate i,

$$S_I = -k_B \sum_i P_i \ln P_i \tag{9.2}$$

which is a more general expression, as it works for any probability distribution. We verify this by considering it for the equal distribution, in which $P_i = \frac{1}{\Omega}$. Then,

$$S = -k_B \sum_{i=1}^{\Omega} = \frac{1}{\Omega} \ln \frac{1}{\Omega} = -k_B \Omega \frac{1}{\Omega} - \ln \Omega = k_B \ln \Omega \tag{9.3}$$

as expected.

We don't want to count the microstates that end up not conserving energy (we can give those zero probability). So, we have these two constraints:

$$\langle E \rangle = \sum_i P_i E_i = U \tag{9.4}$$

$$\sum_i P_i = 1 \tag{9.5}$$

Since entropy wants to maximize itself, we can find the equilibrium state by finding the state with the maximum entropy.

$$\frac{\partial S}{\partial P_i} = 0 \tag{9.6}$$

Using Lagrangian multipliers, we get

$$\frac{\partial}{\partial P_i}\left(\frac{S}{k_B} - \alpha(\text{constraint 1}) - \beta((constraint2))\right) = 0 \tag{9.7}$$

$$\frac{\partial}{\partial P_i}\left(\sum_i (-P_i \ln P_i - \alpha P_i - \beta P_i E_i)\right) = 0 \tag{9.8}$$

Taking derivatives:

$$0 = -\ln P_i - P_i \frac{1}{P_i} - \alpha - \beta E_i \tag{9.9}$$

$$P_i = e^{-1-\alpha-\beta E_i} = \frac{e^{-\beta E_i}}{z} \tag{9.10}$$

(defining $z = e^{1+\alpha}$)

This looks a lot like the Boltzmann distribution. Applying the constraint that $\sum_i P_i = 1$, and that $\sum_i P_i E_i = U$:

$$1 = \sum_i \frac{e^{-\beta E_i}}{z} \implies z = \sum_i e^{-\beta E_i} \tag{9.11}$$

$$\sum_i P_i E_i = U \implies U = \sum_i \frac{E_i e^{-\beta E_i}}{z} \tag{9.12}$$

$$\beta = \frac{1}{k_B T} \implies P_i = \frac{e^{-\frac{E_i}{k_B T}}}{z} \tag{9.13}$$

9.1 Equipartition theorem

Analogous to how energy is proportional to velocity squared or momentum squared, or a length squared (for potential energy), we can say that energy is proportional to efficiency squared. We can find this by taking a continuous version of the Boltzmann distribution, as a function of v, the velocity of a particle in this microstate (cite?)

$\langle E \rangle = \sum_{-\infty}^{\infty} P(v)E(v)dv = \frac{1}{\beta}\frac{\int_{\infty}^{\infty}\left(\beta\frac{1}{2}mv^2\right)e^{-\beta\frac{1}{2}mv^2}dv}{\int_{-\infty}^{\infty}e^{-\beta\frac{1}{2}mv^2}dv}$

Use the change of variable $x = \sqrt{\frac{m\beta}{2}}v$, and cancel a lot of things to get

$$\langle E \rangle = \frac{1}{\beta}\frac{\int_{-\infty}^{\infty} x^2 e^{-x^2} dx}{\int_{-\infty}^{\infty} e^{-x^2} dx} \tag{9.14}$$

A lot of work in polar coordinates will give us

$$\langle E\rangle = \frac{1}{2\beta} = \frac{k_BT}{2} \tag{9.15}$$

This is the equipartition theorem. This lets us find the average energy of an ideal gas. Let the KE of an ideal gas be E_K, then

$$E_K = \frac{1}{2}mv_x^2 + \frac{1}{2}mv_y^2 + \frac{1}{2}mv_z^2 \tag{9.16}$$

$$\langle E_K\rangle = \left\langle \frac{1}{2}mv_x^2 \right\rangle + \left\langle \frac{1}{2}mv_y^2 \right\rangle + \left\langle \frac{1}{2}mv_z^2 \right\rangle = \frac{k_BT}{2} \times 3 = \frac{3}{2}k_BT \tag{9.17}$$

$$\tag{9.18}$$

9.2 Rotational motion in a diatomic gas

A diatomic gas like O_2 has two rotational degrees of freedom; the third would be aligned along the bond axis, and therefore does not affect anything. So

$$\langle E_R\rangle = \left\langle \frac{L_1^2}{2I_1} \right\rangle + \left\langle \frac{L_2^2}{2I_2} \right\rangle = 2\frac{1}{2}k_BT = k_BT \tag{9.19}$$

$$\langle E\rangle = \langle E_T + E_R\rangle = \frac{3}{2}k_BT + k_BT = \frac{5}{2}k_BT \tag{9.20}$$

9.3 Vibrational motion

Vibration has kinetic and potential energies:

$$E = \frac{1}{2}kr^2 + \frac{1}{2}\mu\dot{r}^2 \; \langle E\rangle = \frac{1}{2}k_BT + \frac{1}{2}k_BT = k_BT \tag{9.21}$$

Therefore, with vibrational motion, the total energy of a diatomic particle is $\frac{7}{2}k_BT$. Taking derivatives gives us the C_v and $C_p$per mole of a diatomic gas:

$$C_v = \left(\frac{\partial E}{\partial T}\right)_V = \frac{f}{2}R \tag{9.22}$$

$$C_p = \frac{f+2}{2}R \tag{9.23}$$

In a solid, each atom has three vibration directions, and each vibration has kinetic and potential components. Therefore $\langle E_N\rangle = 3k_BTN$, $\langle C_v\rangle = 3Nk_B$ and $\langle C_p\rangle = 4Nk_B$.

As the solid melts, this description breaks down, as $v \propto \eta^2$. So this only holds for $k_B T << E_{bind}$. Also, we assumed that parameters could take continuous values, which is not true in quantum mechanics. To deal with these, we need to return to an infinite summation rather than an integral. With infinitely small steps in the summation, the summation tends to the integral. So, the integral equipartition theorem holds for $\frac{\Delta}{k_B T} >> 0$.

Physics 5C: Introductory Thermodynamics and Quantum Mechanics **Fall 2018**

Lecture 10: Finding System Parameters from the Probability Distribution

Lecturer: Feng Wang *25 September* *Aditya Sengupta*

10.1 The Partition Function

We previously covered that the probability distribution of a system governed by statistical mechanics is

$$p = \frac{e^{-\beta E_i}}{\sum_i e^{-\beta E_i}} \tag{10.1}$$

We can define $z = \sum_i e^{-\beta E_i}$, the partition function.

Based on this, all of statistical mechanics can be summarized as follows:

1. Calculate z
2. To obtain any parameter (U, S, C_v, P), follow a set procedure based on the probability distribution.

10.2 Single-Particle Function

In a system in which no particles interact with each other, the partition function can be found by multiplying the partition functions of all the particles. Consider a single particle interacting with a reservoir, with fluctuations. Let the system have two discrete energy levels in which the particle can be; an energy of $\frac{\Delta}{2}$ or of $-\frac{\Delta}{2}$.

$$z = \sum_\alpha e^{-\beta E_\alpha} = e^{-\beta\left(-\frac{\Delta}{2}\right)} + e^{-\beta\left(\frac{\Delta}{2}\right)} = 2\cosh\frac{\beta\Delta}{2} \tag{10.2}$$

For N equally spaced energy levels numbered 0 to $N-1$, with separation Δ between each,

$$z = \sum_{\alpha=0}^{N-1} e^{-\beta(\alpha\Delta)} = \frac{1-e^{-N\beta\Delta}}{1-e^{-\beta\Delta}} \tag{10.3}$$

If $N \to \infty$, this simplifies to

$$z = \frac{1}{1-e^{-\beta\Delta}} \tag{10.4}$$

These systems can be described by modelling them as harmonic oscillators. Using a common result from analysis of harmonic oscillators, in which $E_n = n\hbar\omega + \frac{\hbar\omega}{2}$, we can say

$$z_{\text{harmonic oscillator}} = \sum_{\alpha} e^{-\beta\frac{\hbar\omega}{2}} e^{-\beta(\alpha\hbar\omega)} \tag{10.5}$$

$$z_{\text{harmonic oscillator}} = e^{-\beta\frac{\hbar\omega}{2}} \frac{1}{1-e^{-\beta\hbar\omega}} \tag{10.6}$$

10.3 Obtaining functions of state

10.3.1 Internal energy U

We know that

$$\langle U\rangle = \sum_i E_i P_i = \frac{\sum_i E_i e^{-\beta E_i}}{\sum_i e^{-\beta E_i}} \tag{10.7}$$

$$\langle U\rangle = \frac{-\sum_i \frac{\partial}{\partial\beta} e^{-\beta E_i}}{z} = -\frac{\frac{\partial z}{\partial \beta}}{z} \tag{10.8}$$

$$\langle U\rangle = \frac{d\ln z}{d\beta} \tag{10.9}$$

10.3.2 Entropy S

$$S = -k_B \sum_i P_i \ln P_i = -k_B \sum_i \left(\frac{e^{-\beta E_i}}{z} \ln\left(\frac{e^{-\beta E_i}}{z}\right)\right) \tag{10.10}$$

$$S = \frac{-k_B}{z} \sum_i \left(e^{-\beta E_i}(-\beta E_i) - e^{-\beta E_i}\ln z\right) \tag{10.11}$$

$$S = k_B\left(\beta U + \ln z \sum_i P_i\right) = k_B\left(\beta U + \ln z\right) = \frac{U}{T} + k_B \ln z \tag{10.12}$$

10.3.3 Heat capacity at constant volume C_v

$$C_v = \left(\frac{\partial U}{\partial T}\right)_V = \frac{\partial}{\partial T}\left(-\frac{d\ln z}{d\beta}\right)|_V \tag{10.13}$$

$d\beta$ can be rewritten as $d\frac{1}{k_BT} = \frac{-dT}{k_BT^2}$, so

$$C_v = \frac{\partial}{\partial T}\left(\frac{d\ln z}{\frac{dT}{k_BT^2}}\right) = k_BT\left(z\left(\frac{\partial \ln z}{\partial T}\right) + T\left(\frac{\partial^2 \ln z}{\partial T^2}\right)_V\right) \tag{10.14}$$

10.4 Free Energy / the Helmholz function

We know that $dU = TdS - PdV$, which allows us to write U in terms of natural variables (S, V). However, it may be desirable to use a derivative-based representation of U, in terms of (T, V). This leads to the following definition:

$$F = U - TS \tag{10.15}$$

$$dF = dU - TdS - SdT = TdS - PdV - TdS - SdT = -PdV - SdT \tag{10.16}$$

This is also referred to as the *Helmholz free energy.*

10.5 Gibbs Function

This comes out of wanting to express U in terms of (T, P).

$$G = F + PV \tag{10.17}$$

$$dG = dF + PdV + VdP = -SdT + VdP \tag{10.18}$$

This is also referred to as the *Gibbs free energy.* We can define the partition function in terms of these free energies:

$$F = U - TS = U - T\left(\frac{U}{T} + k_B \ln z\right) = -k_B T \ln z \implies z = e^{-\frac{F}{k_B T}} \tag{10.19}$$

These also allow us to redefine pressure and entropy in terms of the Helmholz free energy:

$$P = -\frac{\partial F}{\partial V}_T = k_B T \left(\frac{\partial \ln z}{\partial V}\right)_T \tag{10.20}$$

$$S = -\frac{\partial F}{\partial T}_V \tag{10.21}$$

10.6 Two-Level System Parameters

We previously derived $z = 2\cosh\frac{\beta\Delta}{2}$ for a two-level system. We can find U based on this:

$$U = -\frac{d\ln z}{d\beta} = -\frac{1}{z}\frac{dz}{d\beta} = -\frac{1}{e^{\beta\Delta/2} + e^{-\beta\Delta/2}}\left(\frac{\Delta}{2}e^{\beta\frac{\Delta}{2}} - \frac{\Delta}{2}e^{-\beta\frac{\Delta}{2}}\right) \tag{10.22}$$

$$U = -\frac{\Delta}{2}\tanh\frac{\beta\Delta}{2} \tag{10.23}$$

and similarly

$$F = -\frac{1}{\beta}\ln Z = -k_B T \ln(2\cosh\frac{\beta\Delta}{2}) \tag{10.24}$$

$$S = \frac{U-F}{T} = -\frac{\Delta}{T}\tanh\frac{\beta\Delta}{2} + k_B \ln\left(2\cosh\frac{\beta\Delta}{2}\right) \tag{10.25}$$

$$C_v = k_B\left(\frac{\beta\Delta}{2}\right)^2 {}^2\left(\frac{\beta\Delta}{2}\right) \tag{10.26}$$

We can plot and interpret all of these.

Physics 5C: Introductory Thermodynamics and Quantum Mechanics **Fall 2018**

Lecture 11: Multiparticulate Systems

Lecturer: Feng Wang *September 27* *Aditya Sengupta*

11.1 Modelling Thermodynamic Systems as Harmonic Oscillators

A physics degree is four years of Gaussian integrals and harmonic oscillators:

$$U = \frac{1}{2}kx^2 \tag{11.1}$$

$$E_h = \frac{1}{2}\hbar\omega + n\hbar\omega \tag{11.2}$$

We can use these to write the partition function,

$$Z = \frac{e^{-\frac{1}{2}\beta\hbar\omega}}{1 - e^{-\beta\hbar\omega}} \tag{11.3}$$

$$\therefore U = -\frac{d\ln Z}{d\beta} = -\frac{1}{Z}\frac{dZ}{d\beta} = \frac{\hbar\omega}{2} + \frac{\hbar\omega}{e^{\beta\hbar\omega} - 1} \tag{11.4}$$

In the limit where $k_BT << \hbar\omega$, this simplifies to just $\frac{\hbar\omega}{2}$. In the limit where $k_BT >> \hbar\omega$, we can Taylor expand ($e^x \approx 1 + x$) to get that

$$\frac{\hbar\omega}{\beta\hbar\omega} = \frac{1}{\beta} = k_BT \tag{11.5}$$

We can also plot other system variables,

$$F = -k_BT\ln Z = \frac{\hbar\omega}{2} + k_BT\ln(1 - e^{-\beta\hbar\omega}) \tag{11.6}$$

$$S = \frac{U - F}{T} = k_B\left(\frac{\beta\hbar\omega}{e^{\beta\hbar\omega} - 1} - \ln(1 - e^{-\beta\hbar\omega})\right) \tag{11.7}$$

$$C_v = \frac{\partial U}{\partial T}_V = k_B(\beta\hbar\omega)\frac{e^{\beta\hbar\omega}}{(e^{\beta\hbar\omega} - 1)^2} \tag{11.8}$$

Based on this model, we can plot different system variables. $\frac{U}{\Delta}$ against $\frac{k_BT}{\Delta}$ looks something like an x^2 curve, starting with a relatively flat slope and rising after $k_BT = \Delta$ to become linear. $\frac{S}{k_B}$ rises with a large slope but flattens out, but without approaching a maximum value. C_v against T

11.2 Multiple-Particle Partition Functions

So far, we have assumed all partition functions deal with a single particle, each with its own independent energy levels. We have assumed no interaction between these particles that would affect their probability distributions. We can make a model for multiparticulate systems based on this. Let a system consist of particles $i, j, k, \ldots$. Then we can say the energy is equal to the sum of individual particulate energies:

$$E_{ijk} = E_i^{(a)} + E_j^{(b)} + E_k^{(c)} + \ldots \tag{11.9}$$

which would make the multi-particle partition function equal to

$$Z = \sum_{\{i,j,k,\ldots\}} e^{-\beta E_{total}} = \sum_{\{i,j,k,\ldots\}} e^{-(E_i+E_j+E_k)} = \sum_i \sum_j \sum_k e^{-\beta E_i} e^{-\beta E_j} e^{-\beta E_k} \tag{11.10}$$

$$Z = Z_i Z_j Z_k (\text{splitting the summations and interpreting as partition functions}) \tag{11.11}$$

$$\ln Z = \sum_{\{m=i,j,k,\ldots\}} \ln Z_m \tag{11.12}$$

This is a good approximation if there is no particle interaction. With N independent particles, for example, $Z = (Z_{one})^N$. Also, for diatomic molecules with translational, vibrational, and rotational components to their energy, we can say $Z_{total} = Z_t Z_v Z_r$.

11.3 Statistical Mechanics of Light

In classical mechanics, light is described by electromagnetic waves. We can describe their thermal energy. Consider a metal cube with side length L. This allows us to define boundary conditions on the cube. In the x-direction, $E_{||} = 0$ at $x = 0$ and $x = L$. If we take $E(x) = \sin(k_x x)$, then we apply the $E(L) = 0$ boundary condition, we find that $k_x L = n\pi$. By symmetry, we can extend this to

$$E(x, y, z) = \sin(k_x x)\sin(k_y y)\sin(k_z z) \tag{11.13}$$

$$k_x = n\frac{\pi}{L}, k_y = m\frac{\pi}{L}, k_z = l\frac{\pi}{L} \tag{11.14}$$

n, m, l define different EM modes. Each mode is a different degree of freedom. Let $\vec{k} = (k_x, k_y, k_z)$, then $\omega = c|k|$. We want to know how many modes there are with frequency less than some ω. We can graph this in 3D space; we get integer-valued (scaled by $\frac{\pi}{L}$) nodes on a cube. We can restrict this by requiring that the distance from the origin to the point on the cube is less than $\frac{\omega}{c}$. Then, we can find the number of modes:

$$N = \frac{\text{Total volume}}{\text{Element volume}} = \frac{\frac{1}{8}\frac{4}{3}\pi\left(\frac{\omega}{c}\right)^3}{\left(\frac{\pi}{L}\right)^3} \tag{11.15}$$

$$N = \frac{\omega^3 L^3}{6\pi^2 c^3} = \frac{\omega^3 V}{6\pi^2 c^3} \tag{11.16}$$

We can double this due to polarization (?).

We can also derive a density of state per unit frequency $\frac{dN}{d\omega} = \frac{\omega^2}{\pi^2 c^3} V = g(\omega)$, and based on this an energy of light:

$$U = \int_0^\infty g(\omega) d\omega \times \langle E_0 \rangle = \int_0^\infty \left(k_B T \frac{\omega^2 V}{\pi^2 c^3} \right) d\omega \tag{11.17}$$

since we know the average energy of a state is $k_B T$ (half of this each due to the electric field and magnetic field). We can express this as the product of a spectral energy density and an energy per unit frequency,

$$U(\omega) = k_B T \frac{\omega^2}{\pi^2 c^3} \tag{11.18}$$

This is Rayleigh-Jeans' law. Note that it scales with increasing ω, meaning that a high frequency would correspond to essentially infinite energy. This was referred to as the Rayleigh-Jeans Catastrophe. Since it's unlikely that infinite energy would actually be produced, the experimental scientists involved tried a lot of different mathematical models till they found one that fit the data they had. Experimentally, they found

$$U(\omega) = \frac{\omega^2}{\pi^2 c^3} \left(\frac{\hbar \omega}{e^{\frac{\hbar \omega}{k_B T}} - 1} \right) \tag{11.19}$$

which is the expression that was derived from the harmonic oscillator model before. We can get this analytically from integrating this expression:

$$U = \int_0^\infty \frac{g(\omega)}{V} d\omega \frac{\hbar \omega}{e^{\beta \hbar \omega} - 1} = \frac{\hbar}{\pi^2 c^3} \int_0^\infty \frac{\omega^3 d\omega}{e^{\beta \hbar \omega} - 1} \tag{11.20}$$

$$U = \frac{\hbar}{\pi^2 c^3} \left(\frac{1}{\hbar \beta} \right)^4 \int_0^\infty \frac{x^3 dx}{e^x - 1} = \frac{\pi^2 k_B^4}{15 c^3 \hbar^3} T^4 = A T^4 \tag{11.21}$$

which is the Stefan-Boltzmann Law.

Physics 5C: Introductory Thermodynamics and Quantum Mechanics **Fall 2018**

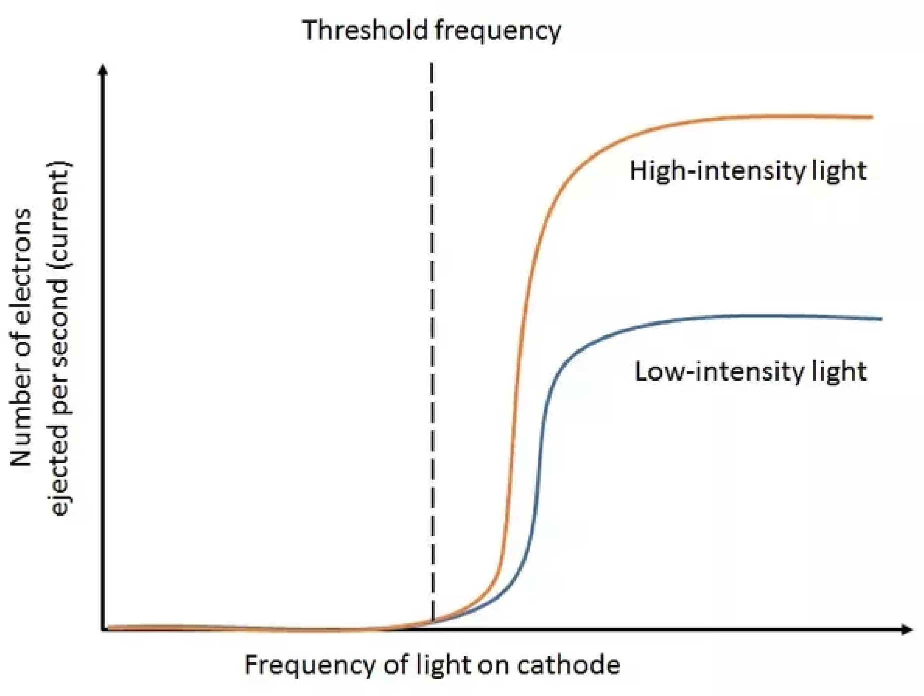
Lecture 12: Quantum Mechanics

Lecturer: Feng Wang *2 October* *Aditya Sengupta*

The theoretical foundation for resolving the Rayleigh-Jeans Catastrophe was Planck's seemingly odd step of assuming energy was quantized, i.e. that

$$E = h\nu \tag{12.1}$$

Based on this, we can define the electron-volt as a microscopic unit of energy: the energy required to move one electron through one volt of potential difference. Thermal energy at 300K is on the order of 25 MeV. A consequence of this quantization is that the photoelectric effect, in which light of a certain freque1ncy causes electronic emission, can be explained; the light is delivering quanta of energy due to their frequency. We can more accurately measure the effect; consider a photoelectric surface at which light is incident at an arbitrary frequency, and a voltage is applied between this and a second plate at voltage $-V_a$. Electrons can cross this barrier if they are given energy at least eV_a. We can measure this by measuring the current in the circuit between the plates, as a function of frequency.



At some critical frequency at which electrons are first moved, $E_k = 0$, after which frequency and E_k scale directly. This effect cannot be purely described by electromagnetism; if it were, then with increased intensity of light at the same frequency, more energy would be delivered and therefore there would be more electronic emission. If we believe that particles can carry quanta of energy based on their frequency, we have an easy explanation for the photoelectric effect; after a work function is overcome at some critical frequency

$h\nu_c = W$, the remaining energy is kinetic: $h(\nu - \nu_c) = E_k$. At greater frequencies, the relationship is not quite linear because the increased photon energy allows for electron excitation and other effects that are explained by convolving density-of-state functions.

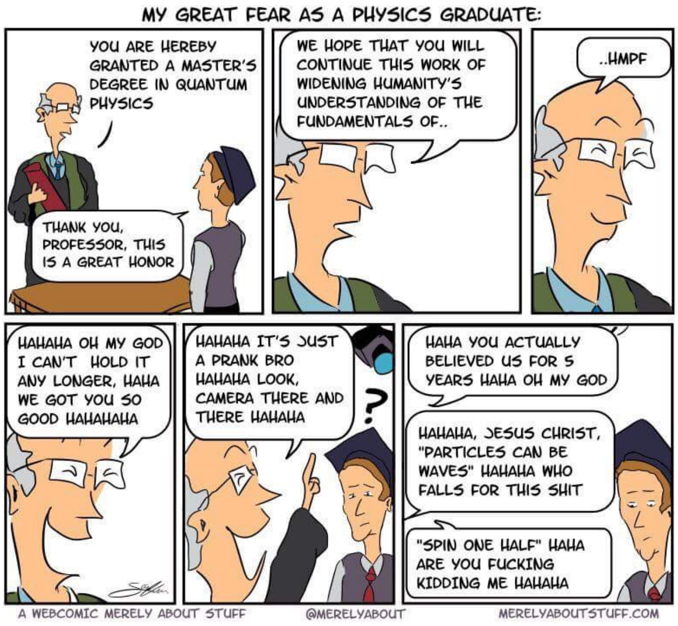
12.1 The Compton effect

For light, $p = \frac{E}{c}$; a definition of momentum allows us to define collisions. An electron with energy E_ν and momentum p_ν can collide with a photon to give it net energy E'_k and to give the electron an energy described by $h\nu'$. In order for the electron and photon masses to be comparable (which must be the case in order for the collision between them to be relatively elastic), we need $m_e \approx m_p = \frac{E}{c^2} = \frac{h\nu}{c^2}$. This happens at the frequency of X-rays.

12.2 Waves or particles?

Describing photons as particles seems to give us nice solutions to a few problems, but we know that light consists of EM waves, which we know how to describe: $E = E_0 e^{i(kx-\omega t)}$. Since both of these explanations seem so useful, we want to reconcile them. A good way to do this would be to explain phenomena that exist in one paradigm in the other one.

Intensity is a simple expression in the particle model, as it scales with energy density, which in turn scales with photon density: $I \propto Nh\nu$. We can replace N with a probability distribution of particles, to connect it with the wave model. Interference can also be explained using the particle model; we can place detectors at both slits in a double-slit experiment in order to detect whether an individual particle is at one point or the other. However, in this case interference will no longer happen, as the photon will always go through one slit and not the other. Observing the state of a system changes the system. This sounds like quantum mechanics proponents are just grasping at straws to keep their theory alive, because honestly, who believes this?



12.3 Atomic structure

We know that an atom consists of protons and electrons. It may seem reasonable to assume that these are homogeneously distributed, as in the Lewis model, but this is wrong, because of the alpha particle experiment; most alpha particles can pass through the atom uninhibited. Therefore the high-mass protons and neutrons are concentrated in the center.

Photon absorption and emission (fill in later)

Balmer series: (the lines - wavelengths of light - being emitted by the hydrogen atom)

$$\lambda = C_2 \left(\frac{n^2}{n^2 - 2^2} \right) \tag{12.2}$$

$C_2 = 0.36456\text{nm}$.

We can take the inverse (multiplying by c) to get the expression for frequency:

$$\nu = \frac{c}{\lambda} = \frac{c}{c_2} \frac{n^2 - 2^2}{n^2} = \frac{4c}{c_2} \left(\frac{1}{2^2} - \frac{1}{n^2} \right) \tag{12.3}$$

Physics 5C: Introductory Thermodynamics and Quantum Mechanics **Fall 2018**

Lecture 13: QM Is Still Weird

Lecturer: Feng Wang *4 October* *Aditya Sengupta*

13.1 Bohr's Model

Experimentally, we have found out that the hydrogen atom has all of its positive charges at the center, that it is stable (i.e. despite electrons accelerating, the system is in equilibrium), and that lines with frequency $\nu = A\left(\frac{1}{2^2} - \frac{1}{n^2}\right)$ are produced. From these conclusions, Bohr was able to make a model of the atom that unified classical and quantum mechanics. The first postulate of Bohr's model is that electrons can have stable orbitals, with discrete energy levels:

$$E_i - E_f = h\nu \tag{13.1}$$

Energy levels are discrete, but we cannot say $E_n \propto n$ for any n, because then a transition between energy states would not produce multiple frequency lines as per the Balmer spectrum. Instead, we can try letting angular momentum take on discrete values.

$$L = n\hbar \tag{13.2}$$

where $\hbar$ is pretty much randomly chosen as a proportionality constant.

In classical mechanics, we know that the Coulombic attraction balances the centripetal force:

$$\frac{e^2}{4\pi\varepsilon_0 r^2} = \frac{mv^2}{r} \tag{13.3}$$

We also know the angular momentum is discretized:

$$L = n\hbar = mvr \tag{13.4}$$

Therefore we can find a radius for the nth orbital:

$$r_n = \frac{4\pi\varepsilon_0 m v^2 r^2}{e^2} = \frac{4\pi\varepsilon_0 (n\hbar)^2}{me^2} = n^2 a_B \tag{13.5}$$

where a_B is the Bohr radius (the radius for $n = 1$).

We can find the energy of the electron,

$$E_n = \frac{1}{2}mv_n^2 - \frac{e^2}{4\pi\varepsilon_0 r_n} = \frac{1}{2}\frac{e^2}{4\pi\varepsilon_0 r_n} - \frac{e^2}{4\pi\varepsilon_0 r_n} \tag{13.6}$$

$$E_n = -\frac{1}{n^2}\frac{e^2}{8\pi\varepsilon_0 a_B} = -\frac{E_0}{n^2} \tag{13.7}$$

Based on this and the energy level difference equation before, we can say that going from lower to higher n corresponds to light absorption, and higher to lower n corresponds to light emission, at certain frequencies. This matches with the hydrogen line emission observations. Look at the energy difference between two states,

$$\Delta E = E_i - E_f = E_0\left(\frac{1}{n_f^2} - \frac{1}{n_i^2}\right) \tag{13.8}$$

which matches with the Balmer line expression exactly as long as $n_f = 2$. Therefore we can see that this model matches up well with real observations.

13.2 Wave Properties

Quantization of angular momentum helps to explain observed properties, but it was not known why energy should be quantized. It was thought that the wave nature of particles could explain this. Enter De Broglie's Hypothesis, in which light could be explained both by properties of electromagnetic waves (λ, ν) and those of photon particles (E, p). We can make relations between these: $E = h\nu$ and $p = \frac{h}{\nu}$. We can connect these expressions via momentum. In the non-relativistic limit, $E_k = \frac{p^2}{2m}$, therefore $p = \sqrt{2m_0E_k}$. We also know that $p = h\lambda$, so

$$\lambda = \frac{h}{\sqrt{2m_0E_k}} \tag{13.9}$$

In general,

$$E_{tot}^2 = E_0^2 + c^2p^2 \tag{13.10}$$

$$p = \frac{1}{c}\sqrt{E_{tot}^2 - E_0^2} \tag{13.11}$$

$$E_0 = m_0c^2; E_{tot} = E_0 + E_k \tag{13.12}$$

$$\therefore \lambda = \frac{hc}{\sqrt{E_{tot}^2 - E_0^2}} = \frac{h/(m_0c)}{\sqrt{\left(\frac{E_k}{E_0}\right)^2 + 2\left(\frac{E_k}{E_0}\right)}} \tag{13.13}$$

We can see that the wavelength scales with $\frac{1}{\sqrt{m}}$.

13.3 Davisson-Germer Experiment

Davisson and Germer essentially scattered light off a crystal and measured its intensity at various angles from 0^o to 90^o, and they got a weird plot, peaking at some angle around 30^o. It is not obvious what this tells us about wave diffraction and interference.

The optical path length difference in light diffracting off crystals is $D \sin\phi$. When this is equal to $n\lambda$, constructive interference is created. At $n = 1$, the wavelength is $2.15 \times sin 50 A = 1.65A$ (get the angstrom symbol).

We can also find the wavelength based on the equation derived from particle properties:

$$\lambda = \frac{h}{(2m_0 e \times 56V)^{1/2}} = 1.66A \tag{13.14}$$

(where 56V is also experimentally determined). Once again, the model follows reality quite well.

Bragg diffraction: $\Delta = 2D\cos\theta = n\lambda$. θ is the complement of ϕ, so they both work. Apparently. Idk.

Physics 5C: Introductory Thermodynamics and Quantum Mechanics **Fall 2018**

Lecture 14: Wave-Particle Duality, the Schrodinger Equation

Lecturer: Feng Wang *9 October* *Aditya Sengupta*

14.1 Thomson's Experiment

The most important consequences of wave-particle duality are that particles can diffract and interfere as if they were waves ($\lambda = D \sin \phi$). This led to Thomson's experiment, in which higher energy electrons were forced to diffract through a thicker layer. This leads to power diffraction, in which the different diffraction peaks make a pattern of concentric circles.

The wavelength of the electron is small, on the order of angstroms. This can be used to our advantage, via electron microscopy. This is a useful technique, but it comes up against the diffraction limit; an electron microscope cannot resolve objects smaller than the wavelength of the electron, about 1 μ m. (Wavelength varies based on energy; an energy of 100 eV corresponds to about 1 angstrom).

14.2 Wave Properties of Atoms and Molecules

It may seem reasonable to say that electrons can behave as waves, but atoms and molecules seem much more tangible. We have working models of ideal gases that depend on their particulate nature. So can we see the wave behaviour of atoms?

Consider the helium atom, the lightest atom that is reasonably stable. We know that

$$\lambda = \frac{h}{\sqrt{2m_0 E_k}} \tag{14.1}$$

For $\lambda = 1A\circ$, $E_k = 100eV$ for the mass of the electron. For the mass of a helium atom, this becomes

$$E_{k,He} = E_{k,E} \times \frac{m_e}{m_{He}} \approx 10meV \tag{14.2}$$

It may seem easier to see the wave nature of the helium atom based on this, as there is less of an energy barrier. However, it is easy to give an electron that kinetic energy by inducing a potential difference (placing it in an electric field). Helium is not so easy to excite, even at this lower energy barrier. Fortunately, thermal energy is on the order of 25 meV, so the wave nature can manifest itself, such as by diffraction once again. Velocities and thermal energies can be selected, and their impact on the diffraction pattern analyzed.

14.3 Wave-Particle Duality

Bohr proposed a resolution to the problem of things interacting both as waves and particles (namely, the problem was that this was weird), called the *complementary principle.* His claim was that the wave and particle pictures are not contradictory, but complementary, and that both are required to describe physical phenomena, but the two cannot be observed simultaneously.

The Born interpretation of this is that photons are described by a wavefunction $\vec{E}(\vec{r}, t)$, from which an intensity $I = |E|^2$ can be observed. From Maxwell's equations, we can derive a wavelength and frequency based on this field. Then, the macroscopic properties of a particle can be described in terms of these variables: $\varepsilon = h\nu = \hbar\omega, p = \frac{h}{\lambda} = \hbar k, \omega = 2\pi\nu, \hbar = \frac{h}{2\pi}, k = \frac{2\pi}{\lambda}$.

The wave and particle aspects cannot be observed at the same time because a wave is distributed over some space, and a particle is localized at a point. The probability of finding the particle at some point x is $\langle x \rangle \propto |E(x)|^2 = I(x)$. Therefore, all particulate properties of an atom that are observed in a discrete sense can be explained, at least mathematically, by extending wave properties.

Electrons and atoms can be completely determined by a wavefunction $\psi(r, t)$, which can be written as a superposition of plane waves with known wave vectors and wavelengths. Anything with spatial dependence or time dependence can be converted to having frequency dependence, which is done by taking a Fourier transform. $E(r, t) = \sum E_n e^{i(k_n \cdot r - \omega_n t)}$. Therefore, the known quantities can be converted to known momenta and energies, which let us derive particulate properties.

14.4 The Schrodinger Equation

Given the initial state of a wavefunction, we want to know how it will evolve over time. This leads us to the Schrodinger equation, which is basically just an educated guess (or an uneducated guess). We need to start from somewhere. Consider plane-wave propagation in the x-direction of photons in free space. We know by Maxwell's equations that this is described by

$$\frac{\partial^2 E}{\partial^2 x} = \frac{1}{c^2}\frac{\partial^2 E}{\partial t^2} \tag{14.3}$$

and in the particulate form, we want to connect this to the particle properties E and p. For a plane wave,

$$E = E_0 e^{-i\omega t + ikx} \tag{14.4}$$

$$\frac{\partial}{\partial t} E_0 e^{-i\omega t + ikx} = (-i\omega) E_0 e^{-i\omega t + ikx} \tag{14.5}$$

$$\frac{\partial E}{\partial x} = ikx E_0 e^{-i\omega t + ikx} \tag{14.6}$$

Therefore a t-derivative can be replaced by multiplication with $-i\omega$, and an x-derivative can be replaced by multiplication with ik. The equation above can therefore be rewritten as

$$(ik)(ik)E = \frac{1}{c^2}(-i\omega)^2 E \tag{14.7}$$

$$\hbar^2 + c^2 k^2 E = \hbar^2 \omega^2 E \tag{14.8}$$

$$c^2 p^2 = \varepsilon^2 \tag{14.9}$$

For electrons, we don't know what the wave description looks like yet, but we know that in particles $E = \frac{p^2}{2m} + V(x)$. Multiply by the wavefunction to get

$$E\psi = \left(\frac{p^2}{2m} + V(x)\right)\psi \tag{14.10}$$

$$\hbar\omega\psi = \left(\frac{\hbar^2 k^2}{2m} + V(x)\right)\psi \tag{14.11}$$

$$\hbar i\frac{\partial\psi}{\partial t} = \left(\frac{\hbar^2}{2m}\left(-i\frac{\partial}{\partial x}\right) + V(x)\right)\psi \tag{14.12}$$

$$i\hbar\frac{\partial\psi}{\partial t} = \left(-\frac{\hbar^2}{2m}\frac{\partial^2}{\partial x^2} + V(x)\right)\psi \tag{14.13}$$

"So this is the Schrodinger equation, based on our educated guesses. In retrospect, it seems so natural." *entire room bursts out laughing.*

"There are many parts of this at which you may ask, why would you do that."

Often, we separate the Schrodinger equation into two parts:

$$i\hbar\frac{\partial\psi}{\partial t} = \hat{H}\psi \tag{14.14}$$

$$\hat{H} = -\frac{\hbar^2}{2m}\frac{\partial^2}{\partial x^2} + V \tag{14.15}$$

where $\hat{H}$ is the Hamiltonian, an operator that acts on a wavefunction.

The rest of this class is basically going to be learning how to do things with the Schrodinger equation. For example, we can characterize wavefunction evolution:

$$\psi(x,0) \to H\psi(x,0) = \left(-\frac{\hbar^2}{2m}\frac{\partial^2}{\partial x^2} + V(x)\right)\psi(x_0) \tag{14.16}$$

$$\psi(x,\Delta t) = \psi(x,0) + \frac{\partial\psi}{\partial t}\Delta t = \psi(x,0) + \Delta t\cdot\frac{1}{i\hbar}\left(H\psi(x,0)\right) \tag{14.17}$$

Essentially, this just repeatedly applies a linear approximation. This is fine for computers, but not for humans.

14.5 Analytical solutions to the Schrodinger equation

If $psi_1, \psi_2, \ldots$ are solutions to the Schrodinger equation, then $\sum_n C_n\psi_n$ is also a solution. Fortunately, we can decompose the energy to get this sort of expansion:

$$E(r,t) = \sum_n E_n(r)e^{-i\omega_n t} \tag{14.18}$$

$$\psi(x,t) = \sum_n \psi_n(x)e^{-i\omega_n t} = \sum_n \psi_n(x)e^{i\frac{E_n}{\hbar}t} \tag{14.19}$$

Each term in this summation is referred to as an *energy eigenfunction.*

Physics 5C: Introductory Thermodynamics and Quantum Mechanics **Fall 2018**

Lecture 15: Understanding and Solving Schrodinger's Equation

Lecturer: Feng Wang *11 October* *Aditya Sengupta*

(not in that order, probably)

15.1 Solving the 1D case (kinda)

Restating Schrodinger's equation in one dimension:

$$i\hbar\frac{\partial}{\partial t}\psi(x,t) = \left(-\frac{\hbar^2}{2m}\frac{\partial^2}{\partial x^2} + V(x)\right)\psi(x,t) = H\psi(x,t) \tag{15.1}$$

To find the general solution, we want to take a superposition of many different linearly independent solutions. These solutions will have well defined energies of the form $E = \hbar\omega$; to convert the wavefunction to frequency space so that we can find this form of energy, we take a Fourier transform:

$$\psi(x,t) = \sum_n \psi_n(x)e^{-i\omega_n t} \tag{15.2}$$

where $\psi_n(x)e^{-i\omega_n t}$ is a state with fixed frequency, or fixed energy. It is often referred to as an *energy eigenstate*. This gives us

$$i\hbar\frac{\partial}{\partial t}\left(\psi_n(x)e^{-i\omega_n t}\right) = \hbar\omega\psi_n(x)e^{-i\omega_n t} = \left(\frac{\hbar^2}{2m}\frac{\partial^2}{\partial x^2} + V(x)\right)\psi_n(x)e^{-i\omega t} \tag{15.3}$$

$$E\psi_n(x) = \hbar\omega\psi_n(x) = \left(\frac{\hbar^2}{2m}\frac{\partial^2}{\partial x^2} + V(x)\right)\psi(x) \tag{15.4}$$

$$\tag{15.5}$$

We've managed to change the Schrodinger equation into a differential equation only in space. Solving this for ψ_n allows us to characterize the time evolution of a one-state wavefunction:

$$\psi(x,t) = \psi_n(x)e^{-i\frac{E_n}{\hbar}t} \tag{15.6}$$

and the probability distribution of the particle:

$$p(x,t) = |\psi(x,t)|^2 = \left|\psi_n(x)e^{-i\frac{E_n}{\hbar}t}\right|^2 = |\psi_n(x)|^2 \tag{15.7}$$

This tells us that the probability distribution is independent of time.

15.2 Is this where cats die? (Spoiler alert: no)

Consider a particle in a box (an infinite square well) whose potential is characterized by

$$V(x) = \begin{cases} 0 & 0 < x < L \\ \infty & \text{otherwise} \end{cases} \tag{15.8}$$

This allows us to split the box into three regions. In the first and third regions, where $x < 0$ or $x > L$ respectively, the potential is infinite therefore there is no chance of finding the particle there. $\psi(x) = 0$. In the second region, we get

$$-\frac{\hbar^2}{2m}\frac{\partial^2}{\partial x^2}\psi_n(x) = E_n\psi_n(x) \tag{15.9}$$

$$\frac{\partial^2}{\partial x^2}\psi_n(x) = -\frac{2mE_n}{\hbar^2}\psi_n(x) = -k_n^2\psi_n(x) \tag{15.10}$$

where k_n is a constant defined for convenience. This has the general solution

$$\psi_n(x) = C_1 \sin k_n x + C_2 \cos k_n x \tag{15.11}$$

We can refine our solution with boundary conditions. One is that $\psi(x = 0) = \psi(x = L) = 0$ for continuity of the wavefunction. At $x = 0$,

$$\psi_n(0) = C_1 \cdot 0 + C_2 \cdot 1 = C_2 = 0 \tag{15.12}$$

So the solution can be simplified to $\psi_n(x) = C_1 \sin k_n x$.

At $x = L$,

$$\psi_n(L) = C_1 \sin k_n L = 0 \tag{15.13}$$

$$C_1 = 0 \text{or} \sin k_n L = 0 \tag{15.14}$$

If C_1 were 0, then the wavefunction would be 0 everywhere. This is technically valid but physically useless. Therefore

$$\sin k_n L = 0 \tag{15.15}$$

$$k_n = \frac{n\pi}{L} \tag{15.16}$$

$$E_n = \frac{n^2\hbar^2\pi^2}{2mL^2} \tag{15.17}$$

We can further restrict this by normalizing the probability distribution:

$$\int_{-\infty}^{\infty} P(x)dx = 1 \tag{15.18}$$

$$1 = \int_0^L \left(C_1 \sin \frac{n\pi x}{L}\right)^2 dx \implies C_1 = \sqrt{\frac{2}{L}} \tag{15.19}$$

In classical mechanics, if we have a bound particle with energy E_n, then $P(x)$ is constant, which is what we see when n becomes very large. So, quantum mechanics is consistent with classical mechanics (at least in this case) as the discretization matters less and less with the availability of more and more eigenstates.

15.3 Superpositions

Let ψ_1, ψ_2 be defined as follows:

$$\psi_1(x) = \sqrt{\frac{2}{L}} \sin \frac{\pi x}{L}, \psi_2(x) = \sqrt{\frac{2}{L}} \sin \frac{2\pi x}{L} \tag{15.20}$$

These translate to space-and-time-dependent wavefunctions which can be added and renormalized:

$$\psi = \frac{1}{2}\left(\sqrt{\frac{2}{L}} \sin \frac{\pi x}{L} e^{i\frac{E_1}{\hbar}t} + \sqrt{\frac{2}{L}} \sin \frac{2\pi x}{L} e^{i\frac{E_2}{\hbar}t}\right) \tag{15.21}$$

What is the energy of this combined state?

$$\langle E \rangle = P_1 E_1 + P_2 E_2 \tag{15.22}$$

In this case, the two states were weighted equally, therefore $E = \frac{E_1+E_2}{2}$. This is only an expected value; a measurement of energy will always yield E_1 and E_2. (In classical systems, energies become very, very high compared to the separation between two eigenvalues, so even though we'd only measure an eigenvalue in a classical system, it's basically the classical energy; experimental error is much greater than this.) This is the collapse of a wavefunction; once a system is measured, it is perturbed, and we get, say, $\psi = \psi_1$.

Over time, the weighting does not stay the same, however, as both evolve with time:

$$p(x,t) = |\psi(x)|^2 = \frac{1}{2}\left|\psi_1 e^{i\frac{E_1}{\hbar}t} + \psi_2 e^{i\frac{E_2}{\hbar}t}\right|^2 \tag{15.23}$$

$$p(x,t) = \frac{1}{2}\left|\psi_1 + \psi_2 e^{i\frac{E_2-E_1}{\hbar}t}\right|^2 \tag{15.24}$$

Physics 5C: Introductory Thermodynamics and Quantum Mechanics **Fall 2018**

Lecture 16: Wavefunction Interpretation

Lecturer: Feng Wang *18 October* *Aditya Sengupta*

16.1 Arrays start at 1

Energy is quantized according to the following (previously derived) formula:

$$E_n = \frac{n^2\hbar^2\pi^2}{2mL^2} \tag{16.1}$$

Interestingly, the zero-point energy actually corresponds to E_1, not $E = E_0 = 0$.

This energy is responsible for the orbital motion of electrons, confining them in a system: $V = \frac{e^2}{r}$. We can find the energy of the first orbital in a 1-angstrom radius innermost shell (terminology cleanup here):

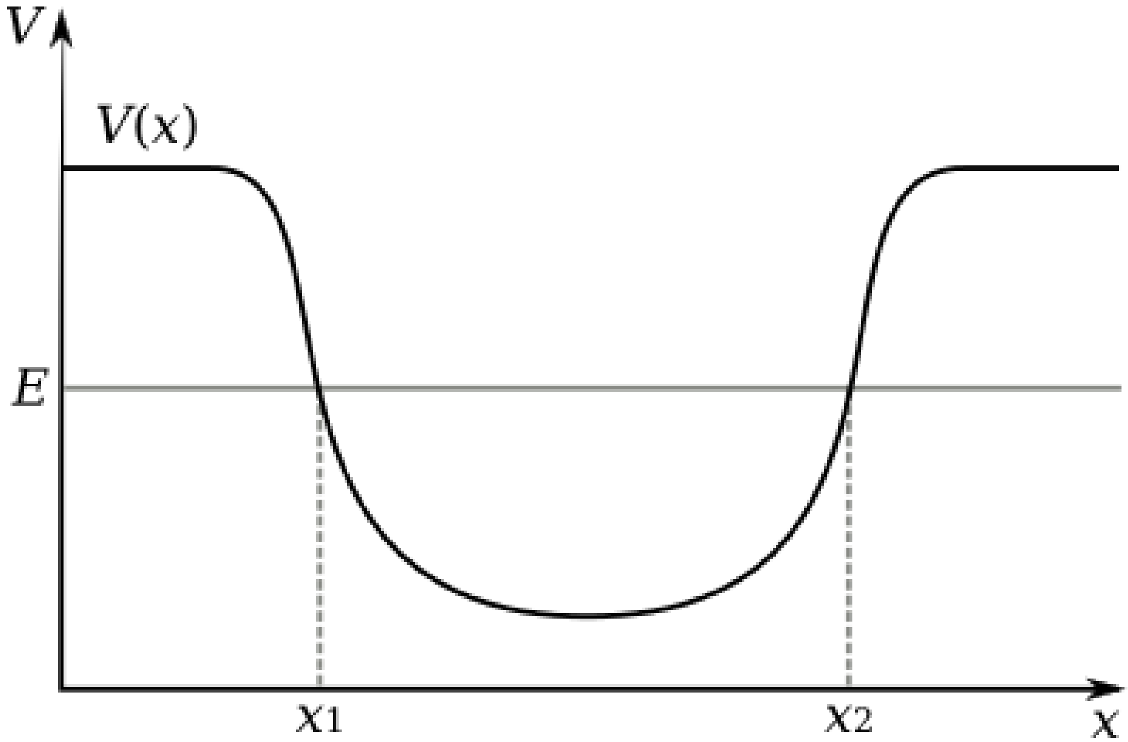
$$\frac{\hbar^2\pi^2}{2m\cdot(1A)^2} = \frac{(\hbar c)^2\pi^2}{2(mc^2)(1A)^2} = \frac{(2000eVA)^2\pi^2}{2\cdot 0.511MeV\cdot A^2} \tag{16.2}$$

which is on the order of 10 eV.

We can analyze the conjecture that the nucleus contains electrons based on this energy expression. The radius of the nucleus is on the order $10^{-4}A$, therefore the energy of the electron becomes $10^9 eV$, which is about a thousand times greater than the mc^2 relativistic energy of the electron. This is physically infeasible, therefore the nucleus cannot contain electrons.

16.2 Particles in a non-rigid box

In a classically-described energy well, certain energies below a threshold after which the slope of the well becomes too high to traverse are "classically allowed", and those above that threshold are "classically forbidden".



Using quantum mechanics, however, we can analyze the *wave* behaviour of a particle that could not classically exist in the forbidden region energy state.

By the Schrodinger equation, we know

$$E\psi = \left(-\frac{\hbar^2}{2m}\frac{\partial^2}{\partial x^2} + V\right)\psi \tag{16.3}$$

$$\frac{\hbar^2}{2m}\frac{\partial^2}{\partial x^2}\psi = (V - E)\psi \tag{16.4}$$

$$\frac{\partial^2\psi}{\partial x^2} = \frac{2m(V - E)}{\hbar^2}\psi = \alpha^2\psi \tag{16.5}$$

The solutions to this are $\psi = e^{\pm\alpha x}$, of which only $e^{-\alpha x}$ makes physical sense. So the wavefunction has overall exponential decay; it peaks in the classically permitted region, but it is nonzero in the classically forbidden region. Quantum mechanically, it can take on classically forbidden energies. The wave that exits the potential well is called an *evanescent wave* (citation needed; all I can find online is E&M related), and this process is called quantum tunneling.

16.3 Qualitative features of bound state wavefunctions

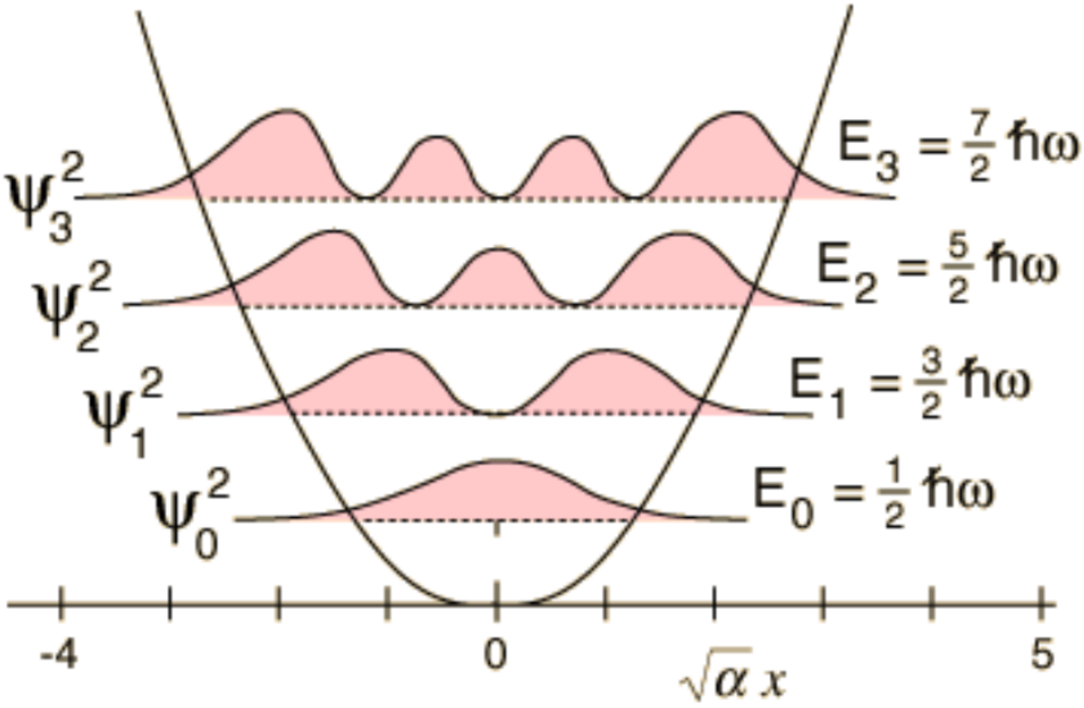
16.3.1 Curvature and oscillatins

We know that

$$\frac{\partial^2\psi}{\partial x^2} = \frac{2m(V-E)}{\hbar^2}\psi \tag{16.6}$$

We can qualitatively analyze this equation to find some trends. If $V > E$, the equation describes exponential decay with positive curvature. If $V >> E$ then the curvature will be very large, so ψ will decrease faster. If $V < E$, the wavefunction will oscillate and have negative curvature. If $V << E$, the oscillations get faster.

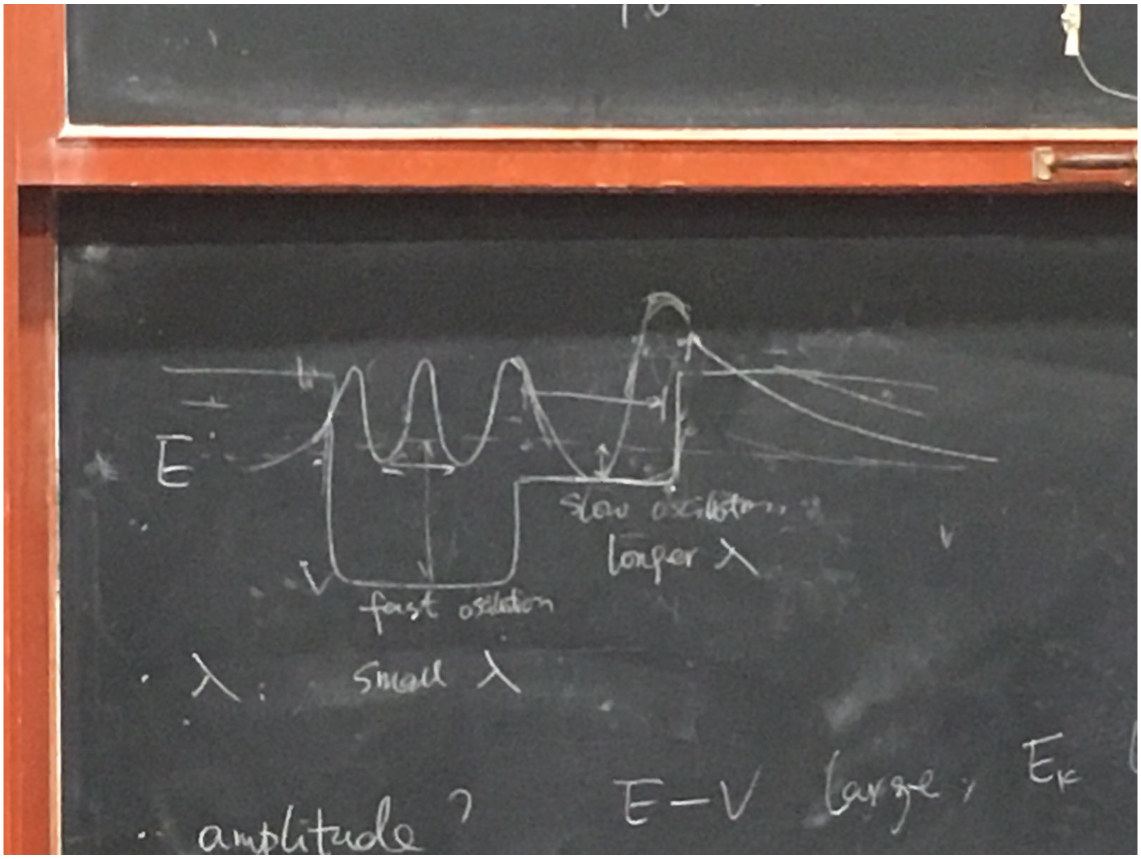
Based on this, what might the wavefunction of a particle with bound energy states look like? At the ground state, $V = 0$, therefore $|0 - E_{min}|$ is the smallest energy level, given that there are many possible energy states in the potential well, which will be oscillating at different rates. Because there are many possible energy states, we may ask whether there is a ground state with no oscillations. To find this, we apply the boundary condition that ψ is continuous and differentiable, i.e. that ψ and $\frac{\partial\psi}{\partial x}$ are both continuous. Otherwise, the second derivative becomes weird to work with. We know that in a bound energy state, the wavefunction has an exponential decay with nonzero, but small and finite, slope. The continuity of the first derivative, coupled with the small and finite slope of the wavefunction at either end, means no energy state can be flat: $E_1 \neq 0$. There must be at least half an oscillation of the wavefunction within the bounding box; this is E_1. Higher and higher energy states are characterized by more oscillations within the box. E_2 has one complete period of a sine wave, and so on; roughly, $\psi_n \propto \sin\frac{nx}{L}$ (I think).



I'm not yet sure why the oscillations in this graphic seem to be in an offset sinc envelope, where the middle oscillations are smaller, rather than all being equal; I'm sure we'll get to that later. For now: higher energies correspond to more oscillations within the bounding box. The ground state has no node (point at which the energy is the averaged energy over the whole box), and after that, the nth eigenstate will have $n-1$ nodes.

16.3.2 Amplitudes

Suppose we have a potential well with two minimum points, first a lower value of V then a higher one. As we saw in the qualitative analysis, the initial larger value of $E - V$ corresponds to faster oscillations and smaller wavelengths (of what?), and the later smaller value of $E - V$ corresponds to slower oscillations and larger wavelengths. How does the amplitude (of the wavefunction?) change? Since large values of $E - V$ correspond to large kinetic energies, v is large, so the particle spends less time in these regions, so the probability is low. Amplitude scales with probability (of finding the particle at a particular point described by the wavefunction), so high energies have low amplitudes.



By counting the number of nodes, we can find which eigenstate of this potential well is represented; in the above image, there are six, so this is the seventh eigenstate.

16.3.3 Symmetry

Symmetry arguments may also be useful in analyzing wavefunctions. If a potential well bottoms out (reaches a minima) at $x = 0$, and has the characteristic $V(x) = V(-x)$, then we can say $|\psi(x)|^2 = |\psi(-x)|^2$. For real-valued $\psi(x)$, we can say that either $\psi(x) = \psi(-x)$ or $\psi(x) = -\psi(-x)$; the first is the general definition for an even wavefunction, and the second is that for an odd wavefunction.

16.3.4 Summary

1. odd/even symmetry for a symmetric potential
2. number of nodes for the nth level is $n-1$
3. relative wavelength (larger for more shallow wells)
4. relative amplitude (larger for more shallow wells)
5. continuity and first x differentiability of ψ

16.4 Quantitatively solving for bound state wavefunctions

Consider a finite square well described piecewise as follows:

$$\begin{cases} V = 0 & -\frac{L}{2} < x < \frac{L}{2} \\ V = V_0 & x > \frac{L}{2} \text{or} x < -\frac{L}{2} \end{cases} \tag{16.7}$$

As previously derived, the wavefunction undergoes exponential growth or decay in the first and third regions (where $V = V_0$), and it oscillates with one frequency in the second region (where $V = 0$).

$$\psi_I = A_1 e^{\alpha x} + A_2 e^{-\alpha x} \tag{16.8}$$
$$\psi_{II} = B_1 \sin kx + B_2 \cos kx \tag{16.9}$$
$$\psi_{III} = C_1 e^{\alpha x} + C_2 e^{-\alpha x} \tag{16.10}$$

We can make use of symmetry arguments to only solve for the first two. Consider the cases of both an even and an odd wave. In an even wave, only the negative exponential and the cosine satisfy the property, so

$$\psi_{I,even} = A_1 e^{\alpha x} \tag{16.11}$$
$$\psi_{II,even} = B_2 \cos kx \tag{16.12}$$
$$\psi_{III,even} = A_1 e^{-\alpha x} \tag{16.13}$$

where we let ψ_I be the positive exponential to account for the fact that x is negative in this region.

To further refine this, we can apply continuity. Take $\psi\left(\frac{L}{2}\right)$:

$$B_2 \cos \frac{kL}{2} = A_1 e^{-\alpha \frac{L}{2}} \tag{16.14}$$
$$\tag{16.15}$$

and take $\psi'\left(\frac{L}{2}\right)$:

$$-kB_2 \sin \frac{kL}{2} = -\alpha A_1 e^{-\frac{\alpha L}{2}} \tag{16.16}$$

We can divide these to get a constraint equation,

$$\tan \frac{kL}{2} = \frac{\alpha}{k} \tag{16.17}$$

Similarly for the odd wave,

$$\psi_{I,odd} = A_1 e^{\alpha x} \tag{16.18}$$
$$\psi_{II,odd} = B_1 \sin kx \tag{16.19}$$
$$\psi_{III,odd} = -A_1 e^{-\alpha x} \tag{16.20}$$

Taking the same boundaries and dividing them in the same way, we get the constraint equation

$$\cot \frac{kL}{2} = -\frac{\alpha}{k} \tag{16.21}$$

Physics 5C: Introductory Thermodynamics and Quantum Mechanics **Fall 2018**

Lecture 17: Fun with the Schrodinger equation

Lecturer: Feng Wang *23 October* *Aditya Sengupta*

17.1 Finite square well wavefunctions

By definition, $k = \frac{(2mE)^{1/2}}{\hbar}$ and $\alpha = \frac{(2m(V_0-E))^{1/2}}{\hbar}$. We can apply this to the constraint equations from before. Since three unknowns are to be determined (E is also an unknown), a third constraint is applied, which is the normalization of the wavefunction; the sum of all probabilities is equal to 1. (Probabilities of finding the particle?)

For the even case, we know that

$$\tan\frac{kL}{2} = \frac{\alpha}{k} \tag{17.1}$$

$$\tan\left(\frac{\sqrt{2mE}}{\hbar}\frac{L}{2}\right) = \left(\frac{V_0}{E} - 1\right)^{1/2} \tag{17.2}$$

and for the odd case,

$$\cot\frac{kL}{2} = -\frac{\alpha}{k} \tag{17.3}$$

$$\cot\left(\frac{\sqrt{2mE}}{\hbar}\frac{L}{2}\right) = -\left(\frac{V_0}{E} - 1\right)^{1/2} \tag{17.4}$$

This is difficult to analytically solve, so we can introduce some graphical methods. First, we make some algebraic replacements:

$$\frac{\sqrt{2mE}}{\hbar}\frac{L}{2} = \theta \tag{17.5}$$

$$\frac{\sqrt{2mV_0}}{\hbar}\frac{L}{2} = \theta_0 \tag{17.6}$$

$$\frac{V_0}{E} = \left(\frac{\theta_0}{\theta}\right)^2 \tag{17.7}$$

Then, for the even case,

$$\tan\theta = \sqrt{\left(\frac{\theta_0}{\theta}\right)^2 - 1} \tag{17.8}$$

and for the odd case,

$$\cot\theta = -\sqrt{\left(\frac{\theta_0}{\theta}\right)^2 - 1} \tag{17.9}$$

We can see on a graph of either case where the decaying square root curve meets the continuously repeating tangent lines, and can numerically estimate the solutions. For the even case, the first solution is close to but less than $\frac{\pi}{2}$, the second is close to but less than $\frac{3\pi}{2}$, and so on incrementing by π in each case until $\theta = \theta_0$. The odd solutions are smaller than but close to integral multiples of π.

In a finite square well, the waves are more spread out than in the infinite case, so the energy in a finite square well is lower. We get this from the graphic solution (how). There is a finite number of bound states in both even and odd wavefunctions. There are more states with larger values of θ_0, leading to a deeper well. There will be at least one bound state (in a one-dimensional case? what's that mean)

17.2 The harmonic oscillator

$$V = \frac{1}{2}cx^2 \tag{17.10}$$

Substituting this into the Schrodinger equation,

$$-\frac{\hbar^2}{2m}\frac{d^2\psi}{dx^2} + \frac{1}{2}cx^2\psi = E\psi \tag{17.11}$$

We examine the limiting behaviour, in which $\psi\ e^{-\alpha x} = e^{-\frac{\sqrt{2m(V_0-E)}}{\hbar}x}$. The harmonic with number 0 acts as

$$\psi\ e^{-\frac{2m(\frac{c}{2}x^2-E)}{\hbar}x} \tag{17.12}$$

which as $x \to \infty$ acts as a Gaussian distribution, $e^{-\frac{\sqrt{mc}}{\hbar}x^2}$. At large x, for the general Gaussian expression, we can take x derivatives:

$$\frac{d\psi}{dx} = \frac{-x}{a^2}\psi \tag{17.13}$$

$$\frac{d^2\psi}{dx^2} = -\frac{1}{a^2}\psi + \frac{x^2}{a^4}\psi \tag{17.14}$$

Therefore, substituting back into the Schrodinger equation gives us

$$-\frac{\hbar^2}{2m}\frac{x^2}{a^4} + \frac{1}{2}cx^2 = 0 \implies a = \left(\frac{\hbar^2}{mc}\right)^{\frac{1}{4}} \tag{17.15}$$

All the eigenstates decay in almost the same way (?) as $x \to \infty$. The ψ_ns are very different, but in general we can say

$$\psi_n(x) = f_n(x) e^{-\frac{x^2}{2a^2}} \tag{17.16}$$

We want to choose an $f_n(x)$ in general such that it is alternating between even and odd. A convenient choice may be $f_n(x) = x^n$. Instead of just large x, we want to ensure this choice fulfils the Schrodinger equation for all x.

$$\psi_0(x) = e^{-\frac{x^2}{2a^2}} \tag{17.17}$$

We just did this one, actually. The only difference is we can no longer set $E = 0$ because we're not in the very-large-x regime, so solving in much the same way as the above yields $E = \frac{ca^2}{2}$.

$$\psi_1(x) = x e^{-\frac{x^2}{2a^2}} \tag{17.18}$$

Substituting this into the Schrodinger equation gives us

$$-a^4 \frac{d^2\psi}{dx^2} + x^2\psi = \frac{2E_1}{c}\psi \tag{17.19}$$

Derivatives suck, but they tell us that

$$E_1 = \frac{3}{2}ca^2 = 3E_0 = 3 \cdot \frac{1}{2}\hbar\omega_0 \tag{17.20}$$

In general, we see that state n has energy $\frac{1}{2}\hbar\omega_0 + (n-1)\hbar\omega_0$.

Physics 5C: Introductory Thermodynamics and Quantum Mechanics **Fall 2018**

Lecture 18: Characterizing Quantum States

Lecturer: Feng Wang *25 October* *Aditya Sengupta*

18.1 Photons

Quantum states are based on observable physical quantities, such as energy, momentum, angular momentum, position. Except that last one; position has to be considered a continuous variable. Momentum can be useful to characterize a physical quantity if you're in free space, but if you have a potential varying as a function of position, it may not be as convenient.

For photons, there's a degree of freedom that's even more convenient, which is somewhat related to angular momentum. We know that for EM waves, in order to define them classically, we have the frequency ω, the propagation direction, and the polarization. An advantage of characterizing a photon by polarization is it exists in two discrete states, perpendicular and parallel, and that it is easy to realize experimentally. It also allows an easy connection to classical EM.

A small disadvantage of using photons is that they can be absorbed and emitted very easily, so the quantity of photons cannot be said to be conserved. A rigorous treatment would require studying quantum electrodynamics, which is a bit much for today. To describe a photon state, it suffices to give its energy, direction, and polarization. For simplicity, we'll mostly consider the polarization.

18.1.1 Linear Polarization

In classical mechanics, the electric field direction determines the polarization completely; if we say propagation is along the $\hat{z}$ direction, then polarization can be in the x-y plane in general. To deal with polarization, we'll use a linear polarizer, which transmits only one electric field direction. If light that is x-polarized passes through a polarizer oriented that way, it'll pass through unchanged; if there is a direction in the middle, it separates it into orthogonal components. In this way, a polarizer can be used as a state analyzer.

Transmission through an arbitrary linear polarizer at an angle θ has intensity described by

$$I = I_0 \cos^2 \theta \tag{18.1}$$

and the part of the electric field in the y direction is $E_y = E_0 \cos \theta$.

We then make a two-part conjecture; that lower energy in the transmitted part corresponding to a lower frequency does not hold for individual photons (is that it?), and that energy is constant with the number of transmitted photons proportional to $\cos^2 \theta$. This doesn't quite make sense if you consider a single photon; we have a probabilistic description of photon transmission or not. Transmission is either 0 or 1, but with a probability given by $\cos^2 \theta$.

The polarization can be considered a quantum mechanical superposition of the x' and y' polarizations, but by passing it through a polarizer, the quantum state is being measured. This changes the quantum state, forcing it into the y' polarization. The probability of finding state A in state B is like a geometric projection.

18.1.2 Circular Polarization

In classical mechanics, we might have an expression looking like

$$\vec{E} = \frac{E_0}{\sqrt{2}} \left(\hat{x} \cos(\omega t - kz) \pm \hat{y} \sin(\omega t - kz) \right) \tag{18.2}$$

We can have left and right circular polarization based on the plus or minus. The projection probability is 50%. Make a projection table:

	X	Y	X'	Y'	R	L
X	1	0	$\cos^2\theta$	$\sin^2\theta$	$\frac{1}{2}$	$\frac{1}{2}$
Y	0	1	$\sin^2\theta$	$\cos^2\theta$	$\frac{1}{2}$	$\frac{1}{2}$
X'	$\cos^2\theta$	$\sin^2\theta$	1	0	$\frac{1}{2}$	$\frac{1}{2}$
Y'	$\sin^2\theta$	$\cos^2\theta$	0	1	$\frac{1}{2}$	$\frac{1}{2}$
R	$\frac{1}{2}$	$\frac{1}{2}$	$\frac{1}{2}$	$\frac{1}{2}$	1	0
L	$\frac{1}{2}$	$\frac{1}{2}$	$\frac{1}{2}$	$\frac{1}{2}$	0	1

A quantum state basis requires orthogonality and completeness. Channels from an analyzer form a set of basis states, with these properties; orthogonality means the projection of X to Y or vice versa is 0, and that any state can be decomposed into a linear combination of these channels. This corresponds to $\sum_i P_i = 1$.

18.2 Generalization of Quantum States

To specify a quantum state, we have to say what kind of particle it is (e.g. an electron), its energy, its direction, and its spin (internal angular momentum).

Physics 5C: Introductory Thermodynamics and Quantum Mechanics **Fall 2018**

Lecture 19: Quantum Amplitudes and Other Made-Up Words

Lecturer: Feng Wang *30 October* *Aditya Sengupta*

19.1 Quantum Amplitudes

The probability distribution of a particle is important, but is not sufficient to describe everything. We introduce the amplitude, which is a complex number directly related to the wavefunction.

One of the key ways of examining photon polarization is an analyzer loop. Consider a polarization analyzer which splits polarized light into basis components (using the X/Y basis, the X'/Y' basis, or the L/R basis). To make light interfere with itself, add a second analyzer that makes the direct lines of light collinear. This is a reversed analyzer. Here, the output is the same thing as the input to the first analyzer, which doesn't seem to serve much of a purpose. But doing this helps to verify that we can split light into its polarization basis elements correctly, and that opens the door to all kinds of analysis. These two analyzers form an analyzer loop. On diagrams, an analyzer loop is denoted by a box in which an oval is drawn, and the two basis elements indicated on the two arcs of the oval.

Consider Y-polarized light incident on an X'-Y' analyzer loop, with an angle of $\theta = 30^\circ$ between Y' and Y. We feed the output of this loop into an X-Y analyzer. We can block any arm of any analyzer in this system; say we block the Y' channel on the analyzer loop (block it between the analyzer and the reverse analyzer). The light is forced to be polarized in the X' state with some probability, so an X' polarization is projected onto the X-Y analyzer. The probability of projection is $\cos^2\theta$; for Y to X', this is 60°, so the probability of this transmission is $\left(\frac{1}{2}\right)^2 = \frac{1}{4}$, and for X' to X, this is 30°, so the probability of this transmission is $\left(\frac{\sqrt{3}}{2}\right)^2 = \frac{3}{4}$. Therefore the overall probability of transmission is $\frac{3}{16}$.

Suppose now that we block the X' channel. The light is forced to be polarized in the Y' state; the probability of transmission of Y to Y' is $\frac{3}{4}$, and the probability of transmission Y' to X is $\frac{1}{4}$. Therefore, again, the probability of transmission overall again is $\frac{3}{16}$.

If both channels are open, the probability becomes 0. This is because the loop analyzer now does nothing, so the setup is just measuring the X polarization of Y-polarized light, which is 0 by definition. Why do two nonzero probabilities combine to make a zero probability? This is because of destructive interference; when only one source of light at a time goes through thin slits, there is some intensity at each point, but it is possible for the two sources to destructively add at certain points where there were nonzero intensities before. More concretely, this comes from $I \propto |E_{total}^2| = |E_1 + E_2|^2$; the magnitudes of the electric fields individually may be nonzero and create nonzero intensities, but their vector sum may end up being zero.

In general, for quantum amplitudes, we know:

1. Probability = $|\text{Quantum amplitude}|^2$
2. Quantum amplitude total = $\sum_i (\text{amplitude})_i$.

19.2 Formalism for projection amplitudes

We represent a projection between state i and state f by $\langle f|i\rangle$, for example $\langle Y|Y'\rangle$. Projections are applied in series by multiplying them together, e.g. $\langle X|X'\rangle\langle X|Y\rangle$. This lets us write the quantum amplitude of the previous system,

$$QA = \langle X|X'\rangle\langle X'|Y\rangle + \langle X|Y'\rangle\langle Y'|Y\rangle = 0 \tag{19.1}$$

$$P = |\langle X|X'\rangle|^2 = \cos^2\theta \implies |\langle X|X'\rangle| = \cos\theta \implies \langle X|X'\rangle = \cos\theta e^{i\phi} \tag{19.2}$$

and we can derive the other polarization projections,

$$|\langle X'|Y\rangle| = \left|\cos\left(\frac{\pi}{2}-\theta\right)\right| \tag{19.3}$$

$$|\langle Y'|Y\rangle| = |\cos\theta| \tag{19.4}$$

$$|\langle X|Y'\rangle = \left|\cos\left(\frac{\pi}{2}+\theta\right)\right| \tag{19.5}$$

Choose the convention where everything is real and positive, we can just drop all the absolute value signs above. In general, the cosine of the angle between each pair of arrows is the probability projection of one onto the other.

Physics 5C: Introductory Thermodynamics and Quantum Mechanics **Fall 2018**

Lecture 20: Linear and Circular Polarization Projections

Lecturer: Feng Wang *1 November* *Aditya Sengupta*

20.1 Linear to Circular Probability Projections

Recall

$$\langle X|X'\rangle = \langle Y'|Y\rangle = \cos\theta \tag{20.1}$$
$$\langle X'|Y\rangle = \langle Y|X'\rangle = \sin\theta \tag{20.2}$$
$$\langle X|Y'\rangle = \langle Y'|X\rangle = -\sin\theta \tag{20.3}$$

We can use the same logic to analyze left and right circular polarization. Consider an R-L analyzer loop, consisting of an LR analyzer followed by an inverted LR analyzer. Place this analyzer loop before an X'-Y' analyzer. Then, consider what happens when R is open, when L is open, and when both are open.

When R is open, we get $\langle Y'|R\rangle \langle R|Y'\rangle$. We know that the probability of projecting linearly polarized light to circularly polarized light is $\frac{1}{2}$, that is, $|\langle Y|R\rangle|^2 = |\langle R|Y\rangle^2| = \frac{1}{2}$, and similarly for the left polarization. Therefore we get a probability of $\frac{1}{4}$ in both polarizations, which gives us a quantum amplitude of

$$\langle Y'|L\rangle \langle L|Y\rangle + \langle Y'|R\rangle \langle R|Y\rangle = \cos\theta \tag{20.4}$$

and a transmission probability of $\cos^2\theta$, the amplitude squared.

Also, due to symmetry, all the components are equal:

$$|\langle Y'|R\rangle| = |\langle R|Y\rangle| = |\langle Y|L\rangle| = |\langle L|Y\rangle| = \frac{1}{\sqrt{2}} \tag{20.5}$$

It is not possible for all the amplitudes to be real, so in general we give all of them complex phases:

$$\langle Y'|R\rangle = \frac{1}{\sqrt{2}}e^{i\alpha} \tag{20.6}$$
$$\langle R|Y\rangle = \frac{1}{\sqrt{2}}e^{i\beta} \tag{20.7}$$
$$\langle Y'|L\rangle = \frac{1}{\sqrt{2}}e^{i\gamma} \tag{20.8}$$
$$\langle L|Y\rangle = \frac{1}{\sqrt{2}}e^{i\delta} \tag{20.9}$$

Then,

$$\langle Y'|R\rangle \langle R|Y\rangle + \langle Y'|L\rangle \langle L|Y\rangle = \left| \frac{1}{2} e^{i(\alpha+\beta)} + \frac{1}{2} e^{i(\gamma+\delta)} \right| = \cos\theta \tag{20.10}$$

From the requirement that the amplitude is equal to $\cos\theta$, we get conditions on the phases: $\alpha = \theta, \beta = 0, \gamma = -\theta, \delta = 0$. There are many different solutions; the general solution has these restrictions:

$$\langle Y|R\rangle = \frac{1}{\sqrt{2}} \tag{20.11}$$

$$\langle R|Y'\rangle = \langle Y'|R\rangle^{*} \tag{20.12}$$

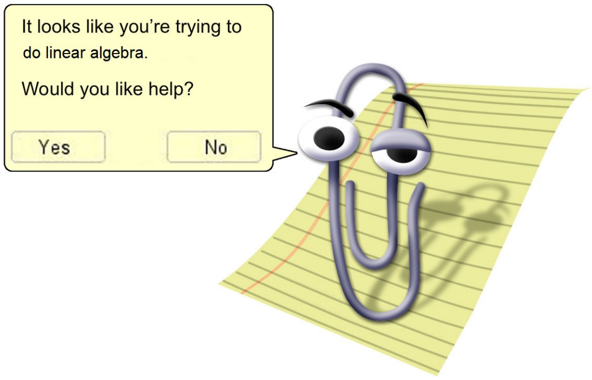
The proof of this is as follows. Consider an arbitrary wavefunction ψ. Its probabilities of left and right polarization have to sum to 1:

$$|\langle R|\psi\rangle|^2 + |\langle L|\psi\rangle|^2 = 1 \implies \langle R|\psi\rangle^{*} \langle R|\psi\rangle + \langle L|\psi\rangle^{*} \langle L|\psi\rangle = 1 \tag{20.13}$$

$$\langle \psi|R\rangle \langle R|\psi\rangle + \langle \psi|L\rangle \langle L|\psi\rangle = 1 \tag{20.14}$$

because the inner product is conjugate-symmetric. We can construct a projection amplitude table,

	—X⟩	—Y⟩	—X'⟩	—Y'⟩	—R⟩	—L⟩
$\langle X\|$	1	0	$\cos\theta$	$-\sin\theta$	$\frac{-1}{\sqrt{2}}$	$\frac{1}{\sqrt{2}}$
$\langle Y\|$	0	1	$\sin\theta$	$\cos\theta$	$\frac{1}{\sqrt{2}}$	$\frac{1}{\sqrt{2}}$
$\langle X'\|$	$\cos\theta$	$\sin\theta$	1	0	$\frac{-1}{\sqrt{2}}e^{i\theta}$	$\frac{1}{\sqrt{2}}e^{-i\theta}$
$\langle Y'\|$	$-\sin\theta$	$\cos\theta$	0	1	$\frac{1}{\sqrt{2}}e^{i\theta}$	$\frac{1}{\sqrt{2}}e^{-i\theta}$
$\langle R\|$	$\frac{1}{\sqrt{2}}$	$\frac{1}{\sqrt{2}}$	$\frac{1}{\sqrt{2}}e^{-i\theta}$	$\frac{1}{2}e^{-i\theta}$	1	0
$\langle L\|$	$\frac{-1}{\sqrt{2}}$	$\frac{1}{\sqrt{2}}$	$\frac{-1}{\sqrt{2}}e^{i\theta}$	$\frac{1}{\sqrt{2}}e^{i\theta}$	0	1



This allows us to express a state in terms of a state vector, with the individual quantum states making up the basis elements being the generalized unit vectors. A vector can be projected to mutually perpendicular axes, $\hat{r} = r_x\hat{x} + r_y\hat{y}$, with $r_x^2 + r_y^2 = 1$.

We can make the projection onto a particular wavefunction into an operator,

$$\langle f|\psi\rangle = \langle f|x\rangle \langle x|\psi\rangle + \langle f|y\rangle \langle y|\psi\rangle |\psi\rangle = \langle x|\psi\rangle |x\rangle + \langle y|\psi\rangle |y\rangle \tag{20.15}$$

This looks suspiciously like a projection onto a basis, $\hat{r} = (\hat{r}\cdot\hat{x})\hat{x}+(\hat{r}\cdot\hat{y})\hat{y}$. $|\psi\rangle$ is a state vector, corresponding to $\hat{r}$, and $|x\rangle$ and $|y\rangle$ form a state vector basis, corresponding to $\hat{x}$ and $\hat{y}$. Then, $\langle x|\psi\rangle$ corresponds to $\hat{r}\cdot\hat{x}$.

This splitting of the bracket into what is functionally a row vector ($\langle Y|$) being multiplied by a column vector ($|Y\rangle$) requires new naming. Since we're splitting the bracket, we might as well split the name "bracket"; the row vector $\langle Y|$ is a "bra" vector, and the column vector $|Y\rangle$ is a "ket" vector.

$$|\psi\rangle = |x\rangle \langle x|\psi\rangle + |y\rangle \langle y|\psi\rangle = |y'\rangle \langle y'|\psi\rangle + |x'\rangle \langle x'|\psi\rangle = |R\rangle \langle R|\psi\rangle + |L\rangle \langle L|\psi\rangle \tag{20.16}$$

A basis for a quantum states must be orthogonal and complete. In bra-ket notation, this comes down to the following conditions on a basis defined with elements i:

$$|\psi\rangle = \sum_i \langle i|\psi\rangle |i\rangle \tag{20.17}$$

$$\sum_i |\langle i|\psi\rangle|^2 = 1 \tag{20.18}$$

20.2 Applying this to the Schrodinger equation

This is all kind of ridiculously abstract. Let's bring it back to wavefunctions. A system has n energy eigenstates, with corresponding wavefunctions $\psi_i(x)$, which can be thought of as a basis for the linear combination of wavefunctions to make a generalized wavefunctions,

$$|\psi\rangle = \sum_i C_i |\psi_i\rangle \tag{20.19}$$

$$\psi(x) = \sum_i C_i \psi_i(x) \tag{20.20}$$

Normalization can give us a condition on the C_is,

$$\int \psi_n^*(x)\psi_n(x)dx = 1 \iff \langle \psi_n|\psi_n\rangle = 1 \tag{20.21}$$

$$\int \psi_n^*(x)\psi_m(x)dx = 0 \iff \langle \psi_m, \psi_n\rangle = 0, m \neq n \tag{20.22}$$

$$C_i = \langle \psi_i|\psi\rangle = \int \psi_i^*(x)\psi(x)dx \tag{20.23}$$

Physics 5C: Introductory Thermodynamics and Quantum Mechanics **Fall 2018**

Lecture 21: State vectors

Lecturer: Feng Wang *6 November* *Aditya Sengupta*

A state vector can be expressed as a linear combination of basis vectors; since the basis vectors are orthonormal, the coefficients are equal to the dot products between the state vector and each basis element:

$$|\psi\rangle = \sum_i |\psi_i\rangle \langle\psi_i|\psi\rangle \tag{21.1}$$

Position representation uses the position state as the basis (invoking the particle nature of the thing), where $|x_i\rangle$ is a position eigenstate. The position states have the projection on one another

$$\langle x_i|x_j\rangle = \delta(x_i - x_j) \tag{21.2}$$

which invokes the Dirac delta.

We can translate between the wavefunction ket and the position representation by projection,

$$|\psi\rangle = \sum_i |x_i\rangle \langle x_i|\psi\rangle \implies \psi(x_i) = \langle x_i|\psi\rangle \tag{21.3}$$

The probability of finding ψ at x_i is therefore

$$|\langle x_i|\psi\rangle|^2 = |\psi(x_i)|^2 \tag{21.4}$$

which turns out to basically be a different formulation of the Schrodinger equation.

21.1 Time dependence of quantum states

Till now we have only discussed particles in a steady state. We can extend this theory to include time dependence as well, starting with a simple case as an example. We previously covered the energy eigenstates to describe the bound states in an infinite square well, which we can revisit to include time dependence,

$$\psi_n(x,t) = \psi_n(x)e^{-i\frac{E_n}{\hbar}t} \tag{21.5}$$

We know that

$$P_n(x,t) = |\psi_n(x,t)|^2 \tag{21.6}$$

which in the stationary state becomes $|\psi_n(x)|^2$. We can construct a superposition of energy eigenstates. Considering the first two energy eigenstates and adding time dependence,

$$\psi(x,t) = A\sin\left(\frac{\pi x}{L}\right)e^{-i\omega_1 t} + A\sin\left(\frac{2\pi x}{L}\right)e^{i\omega_2 t} \tag{21.7}$$

The probability is no longer fixed, as there is a phase change in the time dependent factor (we know that $\omega_2 = 4\omega_1$). We can find the time dependent factor at certain time intervals,

	$e^{-i\omega_1 t}$	$e^{-i\omega_2 t}$
$t = 0$	1	1
$t = \frac{\pi}{\omega_1}$	$e^{-i\omega_1\frac{\pi}{\omega_1}} = -1$	$e^{-i\omega_2\frac{\pi}{\omega_1}} = 1$
$t = \frac{2\pi}{\omega_1}$	1	1

Based on these, we can plot the wavefunction at these fixed times.

Say we have

$$\psi(x,t) = f_1(x)e^{-i\omega_1 t} + f_2(x)e^{-i\omega_2 t} \tag{21.8}$$

$$|\psi(x,t)|^2 = \psi(x,t)\psi^*(x,t) = f_1(x)^2 + f_2(x)^2 + 2f_1(x)f_2(x)\cos(\omega_1 - \omega_2)t \tag{21.9}$$

which is demonstrably independent of the reference zero energy or frequency. Although the particle exists in a well-defined single state, a particle with a superposition of quantum states will have equal probability of an energy measurement showing each of its constituent energy states.

The probability split for the current distribution is

$$|\langle\psi_1|\psi\rangle|^2 + |\langle\psi_2|\psi\rangle|^2 = \left|Ae^{-i(\omega_1 t+\Delta t)}\right|^2 + A^2 = 2A^2 = 1 \tag{21.10}$$

The average energy is $\frac{E_1+E_2}{2}$, although this cannot be measured directly.

To experimentally determine the time-dependent probability distribution of a particle, it is necessary to prepare many copies of the same state, and to do the same measurement at exactly the same time, otherwise interference from the $\cos((\omega_2 - \omega_1)t)$ term would invalidate the equivalence of the states.

Consider the behaviour of a wave packet in a square well. The wavefunction in this state is a superposition of energy eigenstates,

$$\psi(x,t) = \sum_{n=1}^{\infty} B_n \sin\left(\frac{n\pi x}{L}\right)e^{-i\frac{E_n}{\hbar}t} \tag{21.11}$$

At $t = 0$,

$$\psi(x,t) = \sum_{n=1}^{\infty} B_n \sin\left(\frac{n\pi x}{L}\right) \tag{21.12}$$

which is the expression for a Fourier series expansion. We can analyze this further using state orthogonality, the condition that $\langle\psi_j|\psi_i\rangle = 0$. Therefore, in a projection from ψ to ψ_j, we get

$$\langle\psi_j|\psi\rangle = \int \sin\frac{j\pi x}{L}\psi dx \tag{21.13}$$

$$\int \sin\frac{j\pi x}{L}\psi(x,0)dx = \sum_{n=1}^{\infty} B_n \int \sin\frac{n\pi x}{L}\sin\frac{j\pi x}{L}dx = B_j\int_0^L \sin\frac{j\pi x}{L}\sin\frac{j\pi x}{L}dx = B_j\frac{L}{2} \tag{21.14}$$

$$\therefore B_j = \frac{2}{L}\int_0^L \sin\left(\frac{j\pi x}{L}\right)\psi(x,0)dx \tag{21.15}$$

Consider a simple example, with $\psi(x) = A$ where $\frac{L}{2} - \frac{b}{2} < x < \frac{L}{2} + \frac{b}{2}$ and 0 elsewhere. The coefficients are

$$B_n = \frac{2}{L}\int_0^L \sin\frac{n\pi x}{L}\psi(x,0)dx = \frac{2A}{L}\cdot\frac{L}{n\pi}\int_{\frac{L}{2}-\frac{b}{2}}^{\frac{L}{2}+\frac{b}{2}} \tag{21.16}$$

which can, through a LOT of algebra, be found:

$$B_n = \frac{2Ab}{L}\frac{\sin\frac{n\pi b}{2L}}{\frac{n\pi b}{2L}}\sin\frac{n\pi}{2} \tag{21.17}$$

Physics 5C: Introductory Thermodynamics and Quantum Mechanics **Fall 2018**

Lecture 22: Heisenberg Uncertainty, Wave Packet Representations

Lecturer: Feng Wang *8 November* *Aditya Sengupta*

As previously derived, a wave packet in an infinite square well has a wavefunction that is an arbitrary odd Fourier series expansion, where the coefficients are

$$B_n = \frac{2Ab}{L}\frac{\sin\theta}{\theta}\sin\frac{n\pi}{2} \tag{22.1}$$

where $\theta = \frac{n\pi b}{2L}$.

Now, consider a free moving particle without a boundary, i.e. let $L \to \infty$ in the previous case. Here, $\frac{n}{L}$ becomes sufficiently small that the wavefunction can be expressed as an integral. Another way to express this same idea is in a basis defined by particle momentum. For a free particle, we know that an energy eigenstate can be translated to a momentum eigenstate by $E = \frac{p^2}{2m}$, therefore the previous expression in terms of energy becomes

$$\psi_p(x,t) \sim e^{i\frac{p}{\hbar}x}e^{-i\frac{E}{\hbar}t} \sim e^{i(kx-\omega t)} \tag{22.2}$$

This is the expression for a plane wave.

In the free-moving case, consider

$$\psi(x) = \begin{cases} A & -\frac{b}{2} < x < \frac{b}{2} \\ 0 & \text{elsewhere} \end{cases} \tag{22.3}$$

Then, a Fourier transform can be carried out to find the momentum eigenstates present. Using the CTFT analysis equation,

$$B_k = \int \psi(x)e^{-ikx}dx \tag{22.4}$$

These B_ks are synthesized using the CTFT synthesis equation,

$$\psi(x,0) = \int_{-\infty}^{\infty} B(k)e^{ikx}dk \tag{22.5}$$

The position and momentum representations of an eigenstate can be more compactly stated in braket notation,

$$\psi(x) = \langle x|\psi\rangle \text{ - position representation} \tag{22.6}$$
$$B(k) = \langle \psi_k|\psi\rangle \text{ - momentum representation} \tag{22.7}$$

Doing an integral over the above expression for $\psi(x)$ tells us that

$$Ab\left(\frac{\sin\frac{kb}{2}}{\frac{kb}{2}}\right) \tag{22.8}$$

In the graph of sinc, we can measure Δk, the width of the central maximum of the sinc. By varying b, we can see different values of Δk and find that it scales with $\frac{2\pi}{b}$, from which we can find

$$\Delta p \Delta x \sim 2\pi\hbar \tag{22.9}$$

This is the Heisenberg uncertainty principle.

Based on the definition of k, and the statement of the uncertainty principle in terms of it, we can say

$$\Delta\left(\frac{2\pi}{\lambda}\right)\Delta x \tag{22.10}$$

The uncertainty in wavelength can be related to the uncertainty in length. Consider a case where it is not known whether a particle's position lies on the Nth period of the plane wave or the $N+1$th. That is, $\lambda = \frac{L}{N}$ or $\lambda = \frac{L}{N+1}$. This gives us

$$\Delta\lambda = \frac{L}{N} - \frac{L}{N+1} = \frac{L^2}{N(N+1)}\frac{1}{L} = \lambda^2\frac{1}{L} \tag{22.11}$$

Consider

$$\psi(x) = \begin{cases} Ae^{ik_0x} & -\frac{b}{2} < x < \frac{b}{2} \\ 0 & \text{elsewhere} \end{cases} \tag{22.12}$$

Then, using CTFT analysis, we get

$$B_k = \int_{-\frac{b}{2}}^{\frac{b}{2}} Ae^{-i(k-k_0)x}dx = \frac{A\sin((k-k_0)b/2)}{(k-k_0)/2} \tag{22.13}$$

In a more general case, we have

$$\psi(x,t) \sim \int B_k e^{i(kx-\frac{E}{\hbar}t)} dk \sim \int B_k e^{i(kx-\frac{\hbar k^2}{2m}t)} dk \tag{22.14}$$

This allows us to define a group velocity of a wave packet. We know that $\frac{\omega}{k} = \frac{\lambda}{T} = v_p$, the phase velocity. We also know that $\omega = \frac{\hbar k^2}{2m}$, so

$$\frac{d\omega}{dk} = \frac{\hbar k}{m} = \frac{p}{m} = v_{\text{group}} \tag{22.15}$$

Consider a specific case where k is restricted to the range $k_0 \pm \Delta k$. This makes the general integral for ψ equal to

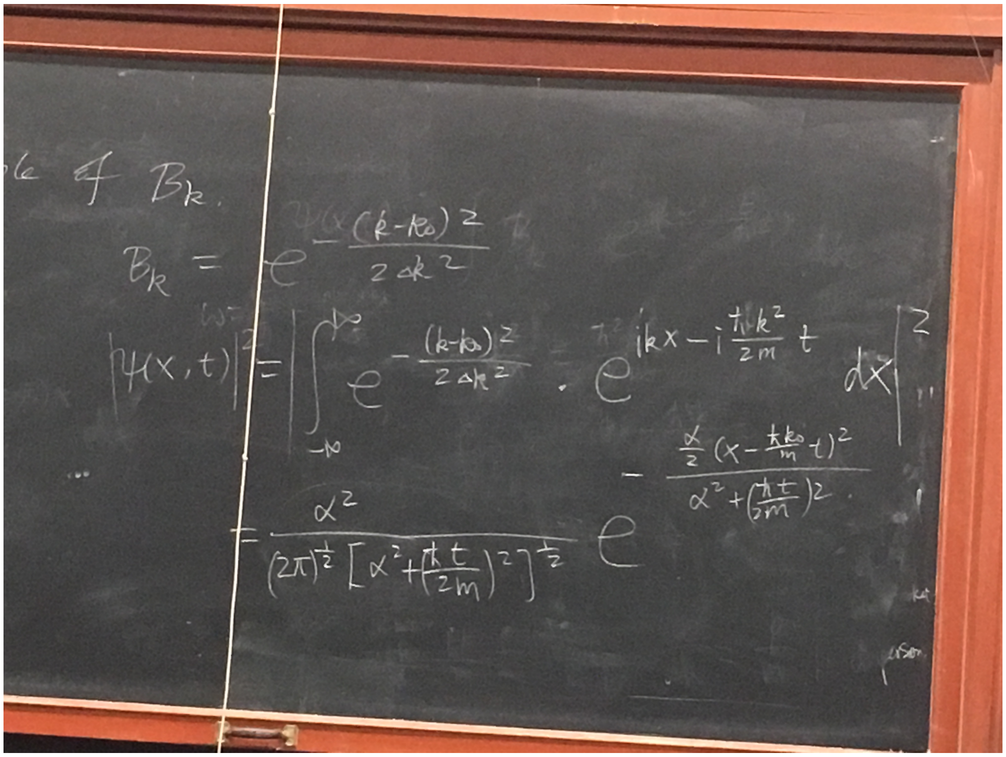
$$\int_{k'=-\Delta k}^{\Delta k} e^{i(k_0+k')x-\frac{\hbar}{2m}(k_0+k')^2 t} dk \tag{22.16}$$

This is an ugly integral, so assuming $\Delta k << k_0$, we get

$$\psi = e^{i(k_0 x - \frac{\hbar k_0^2 t}{2m})} \int_{-\Delta k}^{\Delta k} e^{i\left(x-\frac{\hbar k_0}{m}\right)k'} dk' \tag{22.17}$$

which integrates to something kind of ugly. We can take the magnitude and it looks like a sinc,

$$|\psi(x,t)|^2 = \left| \frac{\sin\left(\left((x-\frac{\hbar k_0}{m}t)\Delta k\right)\right)}{(x-\frac{\hbar k_0}{m}t)\Delta k} \right|^2 \tag{22.18}$$



le of B_k.

$$B_k = e^{-\frac{(k-k_0)^2}{2\Delta k^2}}$$

$$|\psi(x,t)|^2 = \left| \int_{-\infty}^{\infty} e^{-\frac{(k-k_0)^2}{2\Delta k^2}} \cdot e^{ikx - i\frac{\hbar k^2}{2m}t} \, dx \right|^2$$

...

$$= \frac{\alpha^2}{(2\pi)^{\frac{1}{2}}\left[\alpha^2+\left(\frac{\hbar t}{2m}\right)^2\right]^{\frac{1}{2}}} e^{-\frac{\frac{\alpha}{2}\left(x-\frac{\hbar k_0}{m}t\right)^2}{\alpha^2+\left(\frac{\hbar t}{2m}\right)^2}}$$

This looks terrible.

Physics 5C: Introductory Thermodynamics and Quantum Mechanics **Fall 2018**

Lecture 23: Analysis of Uncertainty

Lecturer: Feng Wang *November 13* *Aditya Sengupta*

It is impossible to accurately measure both position and momentum; there is always some spread in position, which can be Fourier transformed to find the uncerainty in momentum. The condition is $\Delta x \cdot \Delta p \geq \hbar$. In measurements, this comes from counting uncertainty, as derived in the previous lecture. We found that $\Delta\lambda \geq \frac{\lambda^2}{\Delta L}$. Defining $\Delta\lambda/\lambda^2 = \Delta k$, we can restate the uncertainty principle as $\Delta k \cdot \Delta l \geq 1$.

The time period of a wave $T = \frac{\Delta\tau}{N}$, where $\Delta\tau$ is the width of a larger wave packet, can be used to find the time-domain uncertainty:

$$\Delta T = \left(\frac{\Delta\tau}{N} - \frac{\Delta\tau}{N+1}\right) \approx \frac{\Delta\tau \cdot \Delta\tau}{N^2}\frac{1}{\Delta\tau} \tag{23.1}$$

$$\frac{\Delta T}{T^2} \cdot \Delta\tau \geq 1 \implies \Delta\omega \cdot \Delta\tau \geq 1 \tag{23.2}$$

Time-frequency uncertainty. Worlds are colliding. Since energy and frequency are directly proportional with proportionality constant $\hbar$, therefore $\Delta E \cdot \Delta\tau \geq \hbar$.

23.1 Consequences of Uncertainty

The energy-time uncertainty means finding the exact energy of a particle would take infinite time. More concrete examples of this uncertainty can be stated. Consider the infinite square well with energies E_1, E_2, E_3. Two electrons in the energy states E_1 and E_2 can be said to form a dipole, which in classical models has decaying oscillations. In quantum mechanics, the excited state decays suddenly according to a probability distribution. Consider the interval from time t to time $t + \Delta t$. The probability of decay ΔP should depend linearly on Δt, a rate constant γ, and on the overall probability P (if the state has decayed the probability of the state decaying in the next time interval is zero). Therefore $\frac{dP}{P} = -\gamma dt \implies P(t) = P_o e^{-\gamma t}$. We know that $P(t) \propto |\psi(t)|^2$, therefore

$$|\psi(t)| = P_o e^{-\frac{t}{\tau}} \tag{23.3}$$

where $\tau = \frac{1}{\gamma}$, the decay time. This characterizes the uncertainty in time of the excited state.

If the atomic transition has a time uncertainty of 10^{-8}s, then the uncertainty in energy turns out to be $\Delta E \sim 10^{-7}$ eV. Something about forbidden transitions.

Consider the case of gamma ray radiation from unstable nuclei. These have a time uncertainty of 10^{-19}s, using the same kind of argument as above.

23.2 Shape of the Spectrum

As derived previously, the wavefunction acts as a decaying exponential, which lets us define amplitude via a Fourier transform,

$$A(E) = \int_{-\inf}^{\inf} \psi(t) e^{i\frac{Et}{\hbar}}\, dt \sim \int_{0}^{\inf e^{-\frac{\gamma}{2}t} e^{i\left(\frac{E-E_0}{\hbar}\right)}} \tag{23.4}$$

$$A(E) = -\frac{1}{\frac{i(E-E_0)}{\hbar} - \frac{\gamma}{2}} \tag{23.5}$$

and the probability as a function of energy is

$$P(E) \sim \frac{1}{(E-E_0)^2 + \left(\frac{\hbar\gamma}{2}\right)^2} \tag{23.6}$$

The full-width half maximum of this distribution is at $E - E_0 = \pm\frac{\hbar\gamma}{2}$, i.e. $\Delta E = \hbar\gamma$. Multiplying by $\Delta t = \frac{1}{\gamma}$ gives us, once again, that $\Delta E \cdot \Delta t = \hbar$.

23.3 Particle Scattering

Scattering is what is most commonly observed from particle experiments. We want to be able to characterize the energy dependence and angular dependence of scattering in 3D. Consider a one-dimensional scattering event; this can be analyzed based on the time-evolution equation of a wave packet, but this is very complicated. A better way would be to consider this as an eigenenergy problem, which allows us to make it a steady-state problem based on plane waves. We want to find the wavefunction for a well-defined energy, which is doable using the Schrodinger equation,

$$E\psi = -\frac{\hbar^2}{2m}\frac{\partial^2}{\partial x^2}\psi + V(x)\psi \tag{23.7}$$

Physics 5C: Introductory Thermodynamics and Quantum Mechanics **Fall 2018**

Lecture 24: Scattering

Lecturer: Feng Wang *15 November* *Aditya Sengupta*

24.1 Scattering by one-dimensional particles

Consider a finite square well obeying $V = V_0 u(x)$. Let a particle have the wavefunction amplitude $A_0 e^{ikx}$, propagating forward until the potential step after which the amplitude is $Be^{ik_2 x}$. At the boundary, particles are scattered and some start moving backwards and have amplitude Ae^{ikx}. For analysis using the Schrodinger equation, this can be split into two regions. In the first region with $V = 0$, we have

$$-\frac{\hbar^2}{2m}\frac{\partial^2}{\partial x^2}\psi(x) = E\psi(x) \tag{24.1}$$

$$\frac{\hbar^2}{2m}k^2 = E \tag{24.2}$$

$$k = \frac{\sqrt{2mE}}{\hbar} \tag{24.3}$$

The wavefunction in this region is therefore

$$\psi = A_0 e^{ikx} + Ae^{-ikx} \tag{24.4}$$

with k defined as above. In the second region, we get

$$\left(-\frac{\hbar^2}{2m}\frac{\partial^2}{\partial x^2} + V_0\right)\psi(x) = E\psi(x) \tag{24.5}$$

$$\frac{\hbar^2 k_2^2}{2m} + V_0 = E \tag{24.6}$$

$$k_2 = \frac{\sqrt{2m(E - V_0}}{\hbar} \tag{24.7}$$

Now, we need to find A and B in terms of A_0 using boundary conditions. We know at $x = 0$, the wavefunction is continuous and its first derivative is also continuous. That is,

$$\psi_I = \psi_{II}|_{x=0} \implies A_0 + A = B \tag{24.8}$$

$$\frac{\partial \psi_I}{\partial x} = \frac{\partial \psi_{II}}{\partial x}|_{x=0} \implies A_0 ik - Aik = Bik_2 \implies A_0 k - Ak = Bk_2 \tag{24.9}$$

$$(A_0 - A)\frac{\sqrt{2mE}}{\hbar} = B\frac{\sqrt{2m(E - V_0}}{\hbar} \tag{24.10}$$

Algebra gives us

$$A = A_0 \frac{k_1 - k_2}{k_1 + k_2} \tag{24.11}$$

$$B = A_0 \frac{2k_1}{k_1 + k_2} \tag{24.12}$$

From this, we can define reflection and transmission coefficients. R is defined below,

$$R = \left|\frac{A}{A_0}\right|^2 = \left(\frac{k_1 - k_2}{k_1 + k_2}\right)^2 \tag{24.13}$$

Consider the case where $E - V_0 << V_0$, i.e. the energy is just above the barrier. In this case, there is little transmission ($R \approx 1$), and $k_2 << k_1$. In the case where $E - V_0 >> V_0$, we get $k_1 \approx k_2$ and $R \approx 0$, so there is little reflection. The reflection coefficient can be graphed as a function of energy.

The transmission coefficient is

$$T = \left|\frac{B}{A_0}\right|^2 \cdot \frac{v_2}{v_1} \tag{24.14}$$

This is slightly different because of the change in potential corresponding to a change in velocity. The transmission coefficient should represent the probability of passing an electron past the potential barrier per unit time. Therefore it should be based on a flux of electrons; the density of electrons that could cross the barrier, multiplied by the ratio of velocities in the old and new potentials. If the electrons have zero velocity, then the transmission probability should be zero.

$$T = \left|\frac{B}{A_0}\right|^2 \cdot \frac{v_2}{v_1} = \left(\frac{2k_1}{k_1 + k_2}\right)^2 \cdot \frac{k_2}{k_1} = \frac{4k_1 k_2}{(k_1 + k_2)^2} \tag{24.15}$$

We can verify this by noting that

$$R + T = \frac{(k_1 - k_2)^2}{(k_1 + k_2)^2} + \frac{4k_1 k_2}{(k_1 + k_2)^2} = 1 \tag{24.16}$$

as it should be.

24.2 Conservation of current

From 5B, we know that changes in local charge density over time come only from an inflow or outflow of current:

$$\frac{d\rho}{dt} = -\nabla \vec{j} = -\frac{\partial j}{\partial x} \tag{24.17}$$

Consider a box from x_0 to $x_0 + \Delta x$, with inflow of current j_{in} and outflow j_{out}. In this case, the total ρ can be found by integrating a constant ρ over that length,

$$\frac{d\int_{x_0}^{x_0+\Delta x} \rho dx}{dt} = j(x_0) - j(x_0 + \Delta x) \tag{24.18}$$

Physically, the charge density comes from the amplitude of the wavefunction (?), therefore we can write

$$\rho = |\psi|^2 = \psi^* \psi \tag{24.19}$$

so on the left hand side of the above expression, we get

$$\frac{d}{dt}\int_{x_0}^{x_0+\Delta x} \psi^*(x,t)\psi(x,t)dx \tag{24.20}$$

Differentiating under the integral gives us

$$\frac{d}{dt}(\psi^*\psi) = \psi^*\frac{d\psi}{dt} + \psi\frac{d\psi^*}{dt} \tag{24.21}$$

Since ψ and ψ^* both satisfy the Schrodinger equation, we know that

$$i\hbar\frac{d\psi}{dt} = \left(-\frac{\hbar^2}{2m}\frac{\partial^2}{\partial x^2} + V(x)\right)\psi \tag{24.22}$$

$$-i\hbar\frac{d\psi^*}{dt} = \left(-\frac{\hbar^2}{2m}\frac{\partial^2}{\partial x^2} + V(x)\right)\psi^* \tag{24.23}$$

Therefore we can use the first time derivatives of ψ and ψ^* in the charge-density expression:

$$\frac{d}{dt}(\psi^*\psi) = \frac{1}{i\hbar}\psi^*\left(-\frac{\hbar^2}{2m}\frac{\partial^2}{\partial x^2} + V(x)\right)\psi - \frac{1}{i\hbar}\psi\left(-\frac{\hbar^2}{2m}\frac{\partial^2}{\partial x^2} + V(x)\right)\psi \tag{24.24}$$

$$\frac{d}{dt}(\psi^*\psi) = \frac{i\hbar}{2m}\left(\psi^*\frac{\partial^2\psi}{\partial x^2} - \psi\frac{\partial^2\psi^*}{\partial x^2}\right) \tag{24.25}$$

Therefore, the integral becomes

$$\int_{x_0}^{x_0+\Delta x} \frac{i\hbar}{2m}\left(\psi^*\frac{\partial^2\psi}{\partial x^2} - \psi\frac{\partial^2\psi^*}{\partial x^2}\right) dx \tag{24.26}$$

Note that using the Fundamental Theorem of Calculus, that is,

$$\int_a^b \frac{\partial F}{\partial x} dx = F|_a^b \tag{24.27}$$

we can write this as the evaluation of $\frac{i\hbar}{2m}\left(\psi^*\frac{\partial\psi}{\partial x} - \psi\frac{\partial\psi^*}{\partial x}\right)$ at $x_0 + \Delta x$, minus that at x_0.

Since the right-hand side is the difference between the evaluation of j at x_0 and $x_0 + \Delta x$, we can set j equal to the negative of the above expression,

$$j \equiv -\frac{i\hbar}{2m}\left(\psi^*\frac{\partial\psi}{\partial x} - \psi\frac{\partial\psi^*}{\partial x}\right) \tag{24.28}$$

For $\psi = Ae^{ikx}$, this expression gives us

$$j = -\frac{i\hbar}{2m} \cdot 2ik \cdot |A|^2 = \frac{\hbar k}{m}|A|^2 = v|A|^2 \tag{24.29}$$

Note that if the wavefunction is real, the charge density is zero as the wavefunction is its own complex conjugate.

24.3 Scattering by a 1D potential well

Consider a potential well in which $V = 0$ for $x < 0$ and $x > L$, and $V = -V_0$ elsewhere. The wavefunction has the form

$$\psi_I = A_0e^{ik_1x} + Ae^{-ik_1x} \tag{24.30}$$

$$\psi_{II} = Be^{ik_2x} + Ce^{-ik_2x} \tag{24.31}$$

$$\psi_{III} = De^{ik_1x} \tag{24.32}$$

Applying continuity of the wavefunction and its derivative at $x = 0$ and $x = L$, and knowing the values of k as follows,

$$k_1 = \frac{\sqrt{2mE}}{\hbar}, k_2 = \frac{\sqrt{2m(E + V_0)}}{\hbar} \tag{24.33}$$

we can calculate A, B, C, and D:

$$k_1(A_0 + A) = (B + C)k_1 \tag{24.34}$$

$$A_0k_1 + A(-k_1) = B(k_2) + C(-k_2) \tag{24.35}$$

$$Be^{ik_2L} + Ce^{-ik_2L} = De^{ik_1L} \tag{24.36}$$

$$k_2Be^{ik_2L} - k_2Ce^{-ik_2L} = k_1De^{ik_1L} \tag{24.37}$$

This gives us a lot of algebra, from which we can calculate the transmission and reflection coefficients,

$$T = \frac{|D|^2 \cdot v_3}{|A_0|^2 \cdot v_1} = \frac{|4k_1k_2|^2}{|(k_2 + k_1)^2e^{-ik_2L} - (k_2 - k_1)^2e^{ik_2L}|^2} \tag{24.38}$$

For $k_1 << k_2$, this is approximately

$$T \approx \frac{16k_1^2k_2^2}{|k_2^2(e^{-ik_2L} - e^{ik_2L})|^2} = \frac{4k_1^2}{k_2^2}\frac{1}{\sin^2(k_2L)} \tag{24.39}$$

which has $T \propto E$.

For $k_1 \approx k_2$, we get

$$T \approx \frac{|4k_1^2|^2}{|(2k)^2e^{-ik_2L}|^2} = 1 \tag{24.40}$$

Interference can also affect transmission. Consider the case where $k_2L = n\pi \implies e^{2ik_2L} = 1$. Then, we get

$$T = \frac{|4k_1k_2|^2}{|(k_2 + k_1)^2 - (k_2 - k_1)^2|^2} = 1 \tag{24.41}$$

This corresponds to perfect transmission.

Physics 5C: Introductory Thermodynamics and Quantum Mechanics **Fall 2018**

Lecture 25: Quantum Tunneling

Lecturer: Feng Wang *27 November* *Aditya Sengupta*

25.1 Tunneling

Send a particle of energy E into a potential barrier of potential V, with $V > E$ and with a square barrier of width L. Classically, the particle can never pass through the potential barrier. However, quantum mechanically, there is a finite probability that it will tunnel through the barrier. This is one of the most counterintuitive results from quantum mechanics - that a particle can exist with finite probability in a classically forbidden region.

Consider a unit-step of potential $V = V_0 u(x)$. We can analyze the behaviour of a particle with energy E using the Schrodinger equation in its time-independent form,

$$E\psi = \left(-\frac{\hbar^2}{2m}\frac{\partial^2}{\partial x^2} + V\right)\psi \tag{25.1}$$

The waves in the first part has the form

$$\psi_I = A_0 e^{ikx} + A e^{-ikx} \tag{25.2}$$

with $k = \frac{\sqrt{2mE}}{\hbar}$.

In the second part, we can say

$$\alpha = \frac{\sqrt{2m(V-E)}}{\hbar} \tag{25.3}$$

and so

$$\psi_{II} = B e^{-\alpha x} \tag{25.4}$$

with this definition of α, because the usual term $E - V$ was flipped and placed under a square root which cancels the i in the usual $ik_1 x$ term in the exponent.

Boundary conditions on the continuity of the wavefunction and its space derivative at $x = 0$ give us

$$A_0 + A = B \tag{25.5}$$

$$ikA_0 + (-ik)A = \alpha B \tag{25.6}$$

and so

$$\frac{A}{A_0} = \frac{ik + \alpha}{ik - \alpha} \tag{25.7}$$

$$\frac{B}{A_0} = \frac{2ik}{ik - \alpha} \tag{25.8}$$

This gives us the reflection coefficient,

$$R = \left| \frac{A}{A_0} \right|^2 = \frac{\alpha^2 + k^2}{\alpha^2 + k^2} = 1 \tag{25.9}$$

There is complete reflection, and no transmission. However, with a finite-width square barrier, this is different. The wavefunction here has three components,

$$\psi_I = A_0 e^{ikx} + Ae^{-ikx} \tag{25.10}$$

$$\psi_{II} = Be^{-\alpha x} + Ce^{\alpha x} \tag{25.11}$$

$$\psi_{III} = De^{ikx} \tag{25.12}$$

There are four boundary conditions here. Continuity of the wavefunction at $x = 0$,

$$A_0 + A = B + C \tag{25.13}$$

Continuity of the derivative at $x = 0$,

$$ikA_0 + (-ik)A = (\alpha)B + \alpha C \tag{25.14}$$

Continuity of the wavefunction at $x = L$,

$$Be^{-\alpha L} + Ce^{\alpha L} = De^{ikL} \tag{25.15}$$

and continuity of the derivative at $x = L$,

$$(-\alpha)Be^{-\alpha L} + \alpha Ce^{\alpha L} = ikDe^{ikL} \tag{25.16}$$

Therefore, algebra gives us

$$\frac{D}{A_0} = \frac{4ik\alpha}{e^{ikL}\left((\alpha+ik)^2 e^{-\alpha L} - (\alpha-ik)^2 e^{\alpha L}\right)} \tag{25.17}$$

The tunneling probability is then given by

$$\frac{v_t|D|^2}{v_i|A_0|^2} = \frac{|D|^2}{|A_0|^2} = \frac{16k^2\alpha^2}{|(\alpha+ik)^2 e^{-\alpha L} - (\alpha-ik)^2 e^{\alpha L}|^2} \tag{25.18}$$

With the assumption that $\alpha L >> 1$, the decaying exponential with that exponent can be violently murdered, and as we hold a towel over its mouth we can write the simplified expression

$$T = \frac{16k^2\alpha^2}{(\alpha^2+k^2)^2 e^{2\alpha L}} = 16\frac{E}{V_0}\left(1 - \frac{E}{V_0}\right) e^{-2\alpha L} \tag{25.19}$$

The exponential decay term tends to dominate here:

$$T \sim e^{-2\alpha L} = e^{-\frac{2}{\hbar}\sqrt{2m(V_0-E)}L} \tag{25.20}$$

With a large mass, this transmission becomes negligible. Even an atom qualifies as "large" here. Consider an electron in a metal trying to tunnel through to another metal with an energy gap. For $E_k \approx 3$eV and $V \approx 10$eV, and a thickness $L \approx 1\mathring{A}$, the exponential term is $1.4 \times 10^{10}\text{m}^{-1} \times 10^{-10}$m, and so the transmission is $e^{-2\times 1.4} \approx \frac{1}{10}$

This is used in the scanning tunneling microscope.

25.2 Probability current

The probability current of the wavefunction in the first region is

$$J_I = v_0|A_0|^2 - v_0|A|^2 \tag{25.21}$$

and in the third is

$$J_{III} = v_0|D|^2 \tag{25.22}$$

Forcing these two to be equal gives $|A_0|^2 - |A|^2 = |D|^2$. In the middle, we can calculate the probability current using the first-principles equation based on the form of the wavefunction,

$$J_{II} = -\frac{i\hbar}{2m}\big((B^*e^{-\alpha x} + C^*e^{\alpha x})(-\alpha Be^{-\alpha x} + C\alpha e^{\alpha x}) - (Be^{-\alpha x} + Ce^{\alpha x})(B^*(-\alpha)e^{-\alpha x} + C^*\alpha e^{\alpha x})\big) \tag{25.23}$$

$$J_{II} = -\frac{i\hbar\alpha}{m}\left(B^*C - BC^*\right) \tag{25.24}$$

For this to be equal to the probability current in the other two regions, B and C have to be finite, and complex with different phases.

25.3 Approximation for barrier penetration

In an arbitrary potential rectangular barrier, we know that

$$\frac{\psi(L)}{\psi(0)} \sim e^{-\frac{2m(V_0-E)^{1/2}}{\hbar}L} \tag{25.25}$$

and

$$\alpha(x) = \sqrt{\frac{2m}{\hbar}(V(x) - E)} \tag{25.26}$$

characterizes the exponential decay. Over a small slice Δx, the ratio of wavefunctions is

$$\frac{\psi(x + \Delta x)}{\psi(x)} \sim e^{-\alpha(x)\Delta x} \tag{25.27}$$

With a linear approximation, we get

$$\psi(x) + \frac{\partial\psi}{\partial x}\Delta x \sim \left(x - \alpha(x)\Delta x\right)\psi(x) \tag{25.28}$$

and therefore

$$\frac{\partial\psi(x)}{\psi(x)} = -\alpha(x)dx \tag{25.29}$$

This gives us a complex exponential in the integral of $-\alpha x$, which is a function of $V(x)$, to characterize the wavefunction ratios and therefore the transmission,

$$T = \left|\frac{\psi(x_2)}{\psi(x_1)}\right|^2 =\sim \exp\left(-2\int_{x_1}^{x_2}\frac{\sqrt{2m(V(x)-E)}}{\hbar}dx\right) \tag{25.30}$$

Physics 5C: Introductory Thermodynamics and Quantum Mechanics **Fall 2018**

Lecture 26: Quantum Entanglement, Information, and Computation

Lecturer: Feng Wang *29 November* *Aditya Sengupta*

26.1 Quantum Entanglement

Quantum entanglement is weird, and we can make weirdness useful. The key concepts that we have seen before related to this are that carrying out a measurement causes a wavefunction to irreversibly collapse, and that quantum states exist in superpositions, like Schrodinger's cat. Classically, we can say that a system has a state $|1\rangle$ or $|0\rangle$, but quantum mechanically, we can say that the state is $\alpha|1\rangle + \beta|0\rangle$. Quantum entanglement is the superposition of a multi-particle state, in which the state of one particle depends on a second particle.

Consider a two-photon system with a total angular momentum of zero. The circular polarization of the two photons has to be equal and opposite for this to be the case. This lets us define the state

$$\psi = |\hbar\rangle_1 |-\hbar\rangle_2 + |-\hbar\rangle_1 |\hbar\rangle_2 \tag{26.1}$$

Experimentally, this is actually quite easy to realize. Consider a pion, which is an unstable particle that can simultaneously emit two photons when it decays. If the pion initially has an angular momentum of zero, then the two photons have to have equal and opposite spin; if one is clockwise, the other must be counterclockwise. These two photons are an entangled state. Note that the polarization is still the same, left or right, for both particles, because they move in opposite directions even though their spin is opposite. The combined state wavefunction can then be written in terms of polarization as

$$\psi_{12} = |R_1\rangle |R_2\rangle + |L_1\rangle |L_2\rangle \tag{26.2}$$

Using the notion of movement in the complex plane, we can write this state wavefunction in terms of an $x - y$ basis:

$$\psi_{12} = |X + iY\rangle_1 |X - iY\rangle_2 + |X - iY\rangle_1 |X + iY\rangle_2 \tag{26.3}$$

which can be expanded,

$$\psi = ((|X\rangle_1 + i|Y\rangle_1)(|X\rangle_2 - i|Y\rangle_2)) + ((|X\rangle_1 - i|Y\rangle_1)(|X\rangle_2 + i|Y\rangle_2)) = |X\rangle_1 |X\rangle_2 + |Y\rangle_1 |Y\rangle_2 \tag{26.4}$$

dropping a normalization factor of $\frac{1}{\sqrt{2}}$ on the complex Cartesian expansion.

This polarization turns out to be just as simple in the $x - y$ basis. What's interesting is when you measure the state of one particle, that collapses the wavefunction and determines the state of the other particle, without interacting with it. This is referred to as the EPR (Einstein-Podolsky-Rosen) paradox. It's referred to as a paradox because it violates locality, the idea that the influence of anything physical should extend only to its local surroundings. More simply, it transmits information faster than the speed of light. It's really okay though, because their initial state is together, so entanglement is a local effect, and although they have encoded information very far away, this reaches two different states *at* the speed of light, so it cannot be used for faster-than-light communication.

26.2 Quantum Information

We can introduce the idea of a quantum bit, or qubit, based on entanglement. Quantum mechanically, in a two-level system, you can have $|0\rangle$ and $|1\rangle$ (spin down and spin up), but you can also have a superposition of the two, $\alpha|1\rangle + \beta|0\rangle$. We can use this for cryptography. It is difficult to send a key (an encoding scheme for encrypted messages) remotely; the current system depends on our computational inability to factorize large numbers, but that may change. However, quantum key distribution - more formally the Bennett-Brassard protocol - is theoretically an unbreakable way of sending a key using photon polarization. Say Alice and Bob want to exchange keys. Alice can send photons to Bob encoding the key. The interesting part of this is that the basis can be freely selected; it can be $|0\rangle, |90\rangle$ or $|-45\rangle, |45\rangle$, for example. Bob can select his basis on his polarizer freely too.

First, Alice sends one of the four states randomly - either of the basis elements from either basis A or B. Say she chooses basis A, and a $|0\rangle$ state. Bob also chooses a basis randomly, by rotating his analyzer. Alice sends a sequence of these, e.g. (for basis elements a, $\bar{a}$ of the basis A). Say Alice's transmissions are

$$a\bar{a}b\bar{a}b\bar{b}\bar{b}a \tag{26.5}$$

and Bob, randomly selecting a basis, reads

$$ab\bar{a}\bar{b}\bar{a}\bar{b}\bar{b}b \tag{26.6}$$

If the bases A and B for Alice and Bob happen to match up, the measurement is reproduced, else Bob reads a random basis element in his basis due to Alice's transmission. Suppose Alice and Bob both publish the bases they use; then Alice and Bob can compare, and can note where their bases match up.These count as correct transmissions. Whenever an eavesdropper intercepts Alice's communications, that measurement irreversibly changes the state, which changes the information Bob receives. Alice and Bob can both communicate in this way and publish some test data; if they agree, they know there is no outside observer.

26.3 Quantum Computation

Consider a basis of spin-up and spin-down ($|1\rangle, |0\rangle$) states. One particle has a possible state of

$$C_1|1\rangle + C_2|0\rangle \tag{26.7}$$

Two particles have

$$C_1 \left|11\right\rangle + C_2 \left|00\right\rangle + C_3 \left|10\right\rangle + C_4 \left|01\right\rangle \tag{26.8}$$

and this grows exponentially; N qubits have 2^N possible states.

Consider an operation U on this state; the superposition state after the operation is

$$C_1 \left\langle U|11\right\rangle + C_2 \left\langle U|00\right\rangle + C_3 \left\langle U|10\right\rangle + C_4 \left\langle U|01\right\rangle \tag{26.9}$$

which has 2^N superposition states, all done in parallel. A measurement still has to be done, collapsing this into only one result even though 2^N states are created. This is a problem for systems design, where many processes run in parallel, and one measurement will yield the relevant result.